CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Towards Efficient 3D Reconstruction from UAV Imagery: Evaluation of Feature Detection, Description and Matching Combinations

Ammar Abbas ELMAS

Computer Engineering Department

February, 2026

CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS APPROVAL

Towards Efficient 3D Reconstruction from UAV Imagery: Evaluation of Feature Detection, Description and Matching Combinations

---

Ammar Abbas ELMAS

Computer Engineering Department

This Doctorate Thesis was evaluated by the following Jury Members on 30/01/2026 and was approved by unanimity of votes.

Jury	: Asst. Prof. Dr. Barış ATA	(Advisor)	…………
	: Prof. Dr. Zekeriya TÜFEKÇİ		…………
	: Assoc. Prof. Dr. Ahmet AYDIN		…………
	: Assoc. Prof. Dr. Çiğdem İNAN ACI		…………
	: Assoc. Prof. Dr. Levent KARACAN		…………

This Thesis was written in the Department of Computer Engineering Institute of Natural and Applied Sciences.

Thesis Number:

Prof. Dr. Sadık DİNÇER
Director
Institute of Natural and Applied Sciences

This work was supported by the TÜBİTAK BİDEB 2211-A National PhD Scholarship Program

CONTENTS

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Towards Efficient 3D Reconstruction from UAV Imagery: Evaluation of Feature Detection, Description and Matching Combinations

Ammar Abbas ELMAS

Advisor: Asst. Prof. Dr. Barış ATA

Department of Computer Engineering

ABSTRACT

This thesis develops a comprehensive efficiency evaluation framework for 3D reconstruction from UAV imagery, addressing the critical need for objective algorithm selection and operational optimization. The research establishes systematic methodologies for evaluating computational efficiency, reconstruction quality, and cross-platform performance through multi-dimensional analysis rather than single-metric assessments. A novel Composite Unsupervised Efficiency Score (CUES) methodology enables multi-dimensional evaluation by integrating four distinct weighting approaches for objective algorithm assessment. The developed framework systematically evaluates 784 feature detection, description, and matching combinations across diverse datasets, providing evidence-based recommendations for algorithm selection, parameter optimization, and timing-based performance metrics. Statistical validation demonstrates consistent methodology reliability with high cross-dataset stability, strong ranking consistency between computing environments, and 95% confidence interval quantification enabling significance testing between algorithm alternatives. The framework includes enhanced COLMAP integration, interactive visualization tools, comprehensive performance databases, and standardized evaluation protocols. Key contributions include parameter optimization guidelines, cross-platform benchmarking across server and mobile architectures, multi-layered synthetic UAV datasets development, and novel efficiency metrics for operational UAV scenarios.

This research establishes the first systematic efficiency evaluation framework for UAV-based 3D reconstruction, providing practical guidance for operational deployment optimization and setting new standards for objective performance assessment in computer vision applications.

This research was supported by TÜBİTAK within the scope of BİDEB 2211-A National PhD Scholarship Program.

Keywords: UAV 3D Reconstruction, Algorithm Evaluation, Feature Extraction, Feature Matching, Efficiency Scoring

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

İHA Görüntülerinden Verimli 3B Yeniden Yapılandırma: Öznitelik Algılayıcı, Tanımlayıcı ve Eşleştirme Kombinasyonlarının Değerlendirilmesi

Ammar Abbas ELMAS

Danışman: Dr. Ögr. Üyesi Barış ATA

Bilgisayar Mühendisliği Anabilim Dalı

ÖZ

Bu tez, İHA görüntülerinden 3B yeniden yapılandırma için kapsamlı bir verimlilik değerlendirme metodolojisi geliştirerek, nesnel algoritma seçimi ve operasyonel optimizasyon konusundaki kritik ihtiyacı ele almaktadır. Araştırma, tek metrik değerlendirmeleri yerine çok boyutlu analiz yoluyla hesaplama verimliliği, yeniden yapılandırma kalitesi ve çapraz platform performansını değerlendirmek için sistematik metodolojiler oluşturmaktadır. Yeni Kompozit Denetimsiz Verimlilik Skoru (CUES) metodolojisi, nesnel algoritma değerlendirmesi için dört farklı ağırlıklandırma yaklaşımını entegre ederek çok boyutlu değerlendirme sağlamaktadır. Bu metodoloji, çeşitli veri kümeleri üzerinde 784 öznitelik algılama, tanımlama ve eşleştirme kombinasyonunu sistematik olarak değerlendirerek algoritma seçimi, parametre optimizasyonu ve zamanlamaya dayalı performans metrikleri için kanıt temelli öneriler sunmaktadır. İstatistiksel doğrulama, veri kümeleri arası yüksek stabilite, farklı bilgisayar mimarileri arasında güçlü sıralama tutarlılığı ve algoritma alternatifleri arasında anlamlılık testine olanak tanıyan %95 güven aralığı nicelendirmesi ile metodoloji güvenilirliğini göstermektedir. Metodoloji, geliştirilmiş COLMAP entegrasyonu, interaktif görselleştirme araçları, kapsamlı performans veri tabanları ve standartlaştırılmış değerlendirme protokolleri içermektedir. Temel katkıları arasında parametre optimizasyon rehberleri, sunucu ve mobil mimariler arasında çapraz platform kıyaslama, çok katmanlı sentetik İHA veri kümeleri geliştirme ve operasyonel İHA senaryoları için yeni verimlilik metrikleri yer almaktadır.

Bu araştırma, İHA tabanlı 3B yeniden yapılandırma için ilk sistematik verimlilik değerlendirme metodolojisini oluşturarak operasyonel uygulama optimizasyonu için pratik rehberlik sağlamakta ve bilgisayarlı görü uygulamalarında nesnel performans değerlendirmesi için yeni standartlar belirlemektedir.

Bu araştırma, BİDEB 2211-A Yurtiçi Doktora Burs Programı kapsamında TÜBİTAK tarafından desteklenmiştir.

Anahtar Kelimeler: İHA 3B Yeniden Yapılandırma, Algoritma Değerlendirmesi, Öznitelik Çıkarma, Öznitelik Eşleştirme, Verimlilik Skorlaması

GENİŞLETİLMİŞ ÖZET

İnsansız Hava Araçları (İHA'lar) günümüzde haritacılık, çevre izleme, tarım, arkeoloji ve şehir planlama gibi çok sayıda alanda veri toplama süreçlerini önemli ölçüde dönüştürmüştür. Esneklikleri, maliyet etkinlikleri ve yüksek çözünürlüklü görüntü yakalama kapasiteleri sayesinde 3B yeniden yapılandırma uygulamaları için uygun platformlar haline gelmişlerdir. Ancak İHA tabanlı 3B yeniden yapılandırma, büyük veri hacimleri, hesaplama karmaşıklığı ve algoritma seçiminde objektif değerlendirme eksikliği gibi önemli zorluklarla karşı karşıyadır. Bu tez, söz konusu zorlukları ele almak için sistematik bir değerlendirme metodolojisi geliştirerek İHA tabanlı 3B yeniden yapılandırma alanında önemli bir boşluğu doldurmaktadır.

Araştırmanın temel katkısı, dört farklı denetimsiz ağırlıklandırma yaklaşımını entegre eden yeni Kompozit Denetimsiz Verimlilik Skoru (CUES) metodolojisinin geliştirilmesidir. Bu metodoloji, bilgi teorisi temelli entropi ağırlıklandırması, çok değişkenli analiz temelli PCA ağırlıklandırması, kontrast ve korelasyon analizi temelli CRITIC ağırlıklandırması ve ayrımcılık gücü ölçümü temelli varyans ağırlıklandırması yaklaşımlarını aritmetik ortalama ile birleştirmektedir. Bu yaklaşım, araştırmacı önyargısını ortadan kaldırırken güçlü teorik temelleri koruyarak objektif algoritma değerlendirmesi sağlamaktadır.

Metodolojinin akademik sağlamlığı, çok katmanlı istatistiksel doğrulama süreci ile kanıtlanmıştır. Yerleşik çok kriterli karar verme teknikleri ile karşılaştırma güçlü korelasyon değerleri göstermiştir. Beş farklı veri kümesi arasında gerçekleştirilen çapraz tutarlılık analizi, metodolojinin farklı operasyonel senaryolar arasında genellenebilirliğini doğrulamıştır. Bootstrap yeniden örnekleme yöntemi kullanılarak hesaplanan %95 güven aralıkları, algoritmaların istatistiksel olarak önemli fark gösterip göstermediğini belirlemek için kullanılmaktadır. Bu detaylı doğrulama süreci, metodolojinin hem teorik sağlamlığını hem de pratik güvenilirliğini göstermektedir.

Araştırma, literatürde kapsamlı düzeyde, OpenCV kütüphanesinde bulunan öznitelik algılayıcıları, tanımlayıcıları ve çeşitli eşleştirme stratejilerinin tüm olası kombinasyonlarını sistematik olarak değerlendirmiştir. Toplam 784 farklı algoritma kombinasyonu, beş çeşitli veri kümesi olan sentetik, Oxford, AirSim, UAV ve drone veri kümeleri üzerinde analiz edilmiştir. Bu veri kümeleri, kontrollü sentetik dönüşümlerden gerçek dünya İHA görüntülerine, standart kıyaslama veri kümelerinden Microsoft AirSim simülasyon ortamında üretilen yeni sentetik veri kümelerine kadar geniş bir spektrumu kapsamaktadır.

Ampirik bulgular, İHA uygulamaları için net algoritma performans hiyerarşilerini ortaya koymuştur. Öznitelik algılayıcı ve tanımlayıcı algoritmalarının sistematik değerlendirmesi sonucunda optimal kombinasyonlar belirlenmiş ve farklı uygulama senaryoları için önerilerde bulunulmuştur. ORB algılayıcısı tüm algılayıcı-tanımlayıcı kombinasyonları arasında 0.7483 ortalama verimlilik skoru ile en yüksek ortalama performansı sergilerken, sentetik ortamlarda 0.8821±0.0371 (%95 güven aralığı) zirve performansa ulaşmaktadır. Tam bina yeniden yapılandırmasında ORB-BEBLID-HAM-BF algılayıcı-tanımlayıcı kombinasyonu 0.8489±0.0095 optimal performans sağlarken, çeşitli İHA senaryolarında GFTT-DAISY-L2-BF algılayıcı-tanımlayıcı kombinasyonu 0.7299±0.0615 skoru ile üstün performans göstermektedir.

Araştırmanın önemli keşiflerinden biri, geleneksel kabulleri sorgulayan bulgulardır. Tipik İHA öznitelik sayıları için kaba kuvvet eşleştirmesinin, yaygın olarak üstün kabul edilen yaklaşık en yakın komşu tabanlı yöntemlerden daha iyi performans gösterdiği tespit edilmiştir. Bu bulgu, alternatif eşleştirme yöntemlerinin avantajlarının yalnızca çok yüksek öznitelik sayılarında ortaya çıktığını, bu sayının ise tipik İHA görüntülerinde nadiren aşıldığını göstermektedir. Ayrıca, uygun mesafe metrikleri ile ikili tanımlayıcıların mobil işlemcilerde özellikle verimli çalıştığı belirlenmiştir.

Çapraz platform değerlendirmesi, modern İHA uygulamalarının artan mobil ve uç bilişim gereksinimleri doğrultusunda kritik öneme sahiptir. Sunucu ve mobil işlemciler arasında gerçekleştirilen analiz, algoritma sıralamalarının 0.994-1.000 aralığında yüksek korelasyon ile tutarlı kaldığını göstermiştir. Sunucu ortamlarından mobil platformlara geçişte minimal performans düşüşü, yüksek performanslı algoritmaların kaynak kısıtlı ortamlarda da uygulanabilirliğini doğrulamıştır. Güven aralığı hesaplamaları, Drone veri kümesinde en dar aralıkları (±0.007-0.010), Sentetik ve UAV veri kümelerinde ise daha geniş aralıkları (±0.06-0.08) ortaya koyarak farklı senaryolardaki performans stabilitesini nicelendirmektedir. Veri kümeleri arası tutarlılık analizi 0.865 ortalama stabilite ile metodolojinin genellenebilirliğini kanıtlamıştır.

Microsoft AirSim simülasyon ortamı kullanılarak özgün bir sentetik veri kümesi geliştirilmiştir. Bu veri kümesi, çok katmanlı uçuş senaryoları, değişken kamera açıları ve kontrollü çevresel koşullar sunarak algoritmaların sistematik değerlendirmesini mümkün kılmıştır. Ayrıca, tek bir görüntüden çeşitli dönüşümler uygulanarak oluşturulan sentetik veri kümesi, algoritmaların spesifik değişimlere karşı sağlamlığının izole edilmiş analizini sağlamıştır. Gerçek dünya İHA görüntüleri, Pix4D projelerinden derlenen çeşitli senaryolar ve standart kıyaslama veri kümeleri ile birlikte, detaylı değerlendirme için güçlü bir temel oluşturmuştur.

COLMAP yazılımı, farklı öznitelik çıkarımı ve eşleştirme kombinasyonlarıyla entegre edilerek 3B model oluşturma yetenekleri genişletilmiş ve nicel karşılaştırmalar yapılmıştır. Bu entegrasyon, çeşitli öznitelik yöntemlerinin 3B yeniden yapılandırma üzerindeki etkisinin derinlemesine incelenmesini sağlamıştır.

Araştırmanın operasyonel katkıları, farklı İHA uygulama senaryoları için kanıt temelli algoritma seçim rehberliği sağlamasıdır. Maksimum verimlilik gerektiren uygulamalar için ORB algılayıcısı tabanlı kombinasyonlar, yüksek doğruluk öncelikli uygulamalar için AKAZE ve SIFT algılayıcısı tabanlı yaklaşımlar, mobil dağıtımlar için optimum kombinasyonlar belirlenmiştir. Bu öneriler, detaylı istatistiksel analiz ve çoklu veri kümesi doğrulaması ile desteklenmektedir.

Metodolojinin genellenebilirliği, farklı veri kümeleri arasında 0.738-0.967 aralığında güçlü tutarlılık göstermesi ile kanıtlanmıştır. Bu tutarlılık, metodolojinin temel algoritma karakteristiklerini yakaladığını ve veri kümesi özelindeki yapaylıklardan etkilenmediğini göstermektedir. Dolayısıyla, kontrollü değerlendirme senaryolarında geliştirilen algoritma seçim kararlarının operasyonel senaryolara etkili şekilde aktarılabileceği doğrulanmıştır.

Araştırmanın akademik etkisi, objektif algoritma değerlendirmesi için yeni bir paradigma oluşturmasıdır. Geleneksel yaklaşımlardan veri odaklı objektif metodolojilere geçiş, İHA tabanlı 3B yeniden yapılandırma topluluğu için kanıt temelli karar verme imkanı sağlamaktadır. Metodolojinin bilgisayarlı görü alanındaki diğer değerlendirme problemlerine uyarlanabilir yapı sunması, araştırmanın etkisini İHA uygulamalarının ötesine taşımaktadır.

Açık bilim ilkelerine uygun olarak, araştırmanın tüm bileşenleri kaynak kodlar, veri kümeleri, değerlendirme araçları ve detaylı sonuç belgeleri dahil olmak üzere herkese açık hale getirilmiştir. Bu yaklaşım, sonuçların tekrarlanabilirliğini sağlarken araştırma topluluğunun bulgular üzerine inşa etmesini mümkün kılmaktadır. Geliştirilmiş Plotly tabanlı interaktif görselleştirme araçları, karmaşık performans ilişkilerini anlaşılabilir hale getirerek araştırmacıların ve uygulayıcıların bilgili algoritma seçim kararları vermelerini desteklemektedir.

Bu tez, İHA tabanlı 3B yeniden yapılandırmada objektif algoritma değerlendirmesi için yeni bir standart oluştururken, detaylı ampirik bulgular ve pratik uygulama rehberliği ile alandaki önemli bir boşluğu doldurmaktadır. Metodolojinin istatistiksel doğrulaması, %95 güven aralığı nicelendirmesi ve çapraz platform tutarlılığı, araştırmanın akademik sağlamlığını ve pratik değerini kanıtlamaktadır. Gelecekteki araştırmalar, bu değerlendirme metodolojisini temel alarak derin öğrenme tabanlı yaklaşımların entegrasyonu, gerçek zamanlı işleme gereksinimlerinin ele alınması ve özelleşmiş İHA uygulama alanlarına adaptasyon gibi yönlerde ilerleyebilecektir.

Doktora Adayı Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) Bilim İnsanı Destekleme Daire Başkanlığı (BİDEB) 2211-A Genel Yurt İçi Doktora Burs Programı kapsamında desteklenmiştir.

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to all who have supported me throughout this doctoral journey.

I would like to extend my sincere gratitude to my supervisor, Dr. Barış ATA, for his consistent guidance and patient direction throughout this work. His academic mentorship was instrumental in shaping this thesis to its final state, and I am grateful for his continuous support.

I also extend my gratitude to the members of my Thesis Committee, Prof. Dr. Zekeriya TÜFEKÇİ and Assoc. Prof. Dr. Ahmet AYDIN, for their thoughtful comments and patience throughout the research process. Furthermore, I would like to thank Assoc. Prof. Dr. Çiğdem İNAN ACI and Assoc. Prof. Dr. Levent KARACAN for accepting to serve as jury members for my thesis defense and for their valuable review.

Beyond the academic world, I extend my heartfelt thanks to my family. To my sweet little troublemaker, Ayça, who came into this world as I started this journey, your presence and smiles brought great happiness from the very beginning. And to my dear wife, Gülnur, your constant support and belief in me, no matter what, made the completion of this thesis possible. Your love and patience were my absolute and perhaps sole strength.

This academic pursuit would not have been possible without the collective support of these individuals.

I would like to express my gratitude to The Scientific and Technological Research Council of Türkiye (TÜBİTAK), specifically the Department of Science Fellowships and Grant Programme (BİDEB), for supporting me economically throughout this doctoral process under the 2211-A National PhD Scholarship Program.

LIST OF TABLES

LIST OF FIGURES

SYMBOLS AND ABBREVIATIONS

3D	: 3 Dimensions
AGAST	: Adaptive and Generic Accelerated Segment Test
AKAZE	: Accelerated-KAZE
BEBLID	: Boosted Efficient Binary Local Image Descriptor
BRIEF	: Binary Robust Independent Elementary Features
BRISK	: Binary Robust Invariant Scalable Keypoints
CNN	: Convolutional Neural Network
DoG	: Difference-of-Gaussians
FAST	: Features from Accelerated Segment Test
FLANN	: Fast Library for Approximate Nearest Neighbors
FREAK	: Fast Retina Keypoint
GFTT	: Good Features to Track
GFTT-H	: GFTT with Harris
HL	: Harris Laplace
İHA	: İnsansız Hava Aracı
IoT	: Internet of Things
LATCH	: Learned Arrangements of Three Patch Codes
LoG	: Laplacian of Gaussian
LUCID	: Locally Uniform Comparison Image Descriptor
MAGSAC	: Marginalizing Sample Consensus
MCDM	: Multi-Criteria Decision-Making
MSD	: Maximal Self-Dissimilarities
MSER	: Maximally Stable Extremal Regions
MVG	: Multi-View Geometry
MVS	: Multi-View Stereo
PCA	: Principal Component Analysis
ORB	: Oriented FAST and Rotated BRIEF
RANSAC	: Random Sample Consensus
SfM	: Structure from Motion
SIFT	: Scale Invariant Feature Transform
SLAM	: Simultaneous Localization And Mapping
TEBLID	: Triplet-based Efficient Binary Local Image Descriptor
TBMR	: Tree-Based Morse Regions
UAV	: Unmanned Aerial Vehicles
VGG	: Visual Geometry Group

1. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly referred to as drones, constitute an increasingly significant technology for data acquisition across sectors including surveying, mapping, and environmental monitoring (Vacca et al., 2017). Their operational flexibility, cost-effectiveness, and capacity for capturing high-resolution imagery render them particularly suitable for 3D reconstruction applications (Jiang et al., 2022), especially where traditional methods such as manned aircraft or satellites prove impractical or prohibitively expensive (Bazi et al., 2014). The development of Structure from Motion (SfM) techniques has expanded UAV applications for 3D reconstruction, enabling creation of detailed and accurate 3D models (Westoby et al., 2012).

Implementation of drone-based 3D reconstruction confronts several challenges (Cao, Jia, Lv, et al., 2019). Large quantities of high-resolution imagery generated by modern UAV platforms, combined with computational complexity of photogrammetric processing algorithms, create obstacles to achieving efficient and timely reconstructions (Bianco et al., 2018). Feature extraction and matching represent the foundational phase within SfM processing pipelines, directly affecting both geometric quality and computational efficiency of subsequent reconstruction operations. Contemporary algorithms demonstrate effectiveness in controlled environments (Cao, Jia, Li, et al., 2019), yet often encounter difficulties when processing UAV imagery characteristics including oblique viewing geometries, scale and resolution variations, environmental noise, atmospheric artifacts, and dynamic lighting conditions (Xu et al., 2014). These factors create complex trade-offs between reconstruction accuracy, computational resource requirements, and processing time constraints that practitioners must navigate without adequate guidance from objective evaluation frameworks.

A significant limitation in current UAV-based 3D reconstruction practices involves the absence of objective methodologies for evaluating and selecting optimal feature detection, description, and matching algorithms. Traditional approaches to algorithm assessment often rely on subjective weighting schemes, researcher-defined importance hierarchies, or single-metric evaluations that fail to capture multi-dimensional algorithmic efficiency in real-world deployment scenarios. This gap creates barriers between theoretical algorithm capabilities described in academic literature and practical deployment decisions required in operational environments, where practitioners must balance computational efficiency, reconstruction accuracy, cross-platform compatibility, and resource constraints without evidence-based guidance (Radoglou-Grammatikis et al., 2020).

This thesis addresses these limitations by developing a novel CUES methodology providing objective, statistically rigorous, and detailed algorithm evaluation specifically designed for UAV-based 3D reconstruction applications. The research systematically investigates performance characteristics of 784 distinct combinations of feature detection, description, and matching techniques available in the widely-adopted OpenCV computer vision library. The thesis establishes and validates a theoretically grounded evaluation framework that integrates four distinct unsupervised weighting approaches: entropy, Criteria Importance Through Intercriteria Correlation (CRITIC), Principal Component Analysis (PCA), and variance. This framework eliminates subjective bias in algorithm selection and provides practical, evidence-based guidance for algorithm deployment decisions across diverse operational scenarios.

1.1. Background

3D reconstruction constitutes a digital process for creating spatial representations of physical objects and environments. Originally developed within specialized research contexts, this technology has expanded into practical applications across numerous domains where accurate spatial understanding, geometric analysis, and virtual representation capabilities provide operational advantages.

Cultural heritage preservation and archaeology utilize detailed 3D models for documentation, digital archiving, virtual museum experiences, and restoration planning for historical sites, artifacts, and architectural structures (Xiang et al., 2019). Environmental monitoring and ecological research employ 3D reconstruction for quantifying landscape changes, analyzing vegetation patterns, monitoring erosion processes, assessing habitat structures, and evaluating ecosystem health across temporal scales (Gaffey & Bhardwaj, 2020).

Urban planning and smart city development utilize 3D models for visualization, analysis, and simulation of urban environments, supporting decision-making processes in infrastructure development, transportation planning, energy optimization, and urban growth management (Crommelinck et al., 2016). Construction and infrastructure management sectors deploy 3D reconstruction for monitoring building progress, conducting safety inspections, identifying structural defects, planning maintenance operations, and documenting as-built conditions with millimeter-level precision (Jiang et al., 2017). Agriculture and forestry applications use 3D reconstruction for precision farming, crop health assessment, yield estimation, forest inventory, biomass calculation, and resource optimization (Seifert et al., 2019).

Development and widespread adoption of UAV technology has transformed spatial data acquisition methodologies, particularly in photogrammetric applications where unmanned platforms provide capabilities for efficient, cost-effective, and flexible data collection (Shahbazi et al., 2015). UAV-based photogrammetry represents a shift from traditional aerial platforms, offering democratized access to high-resolution spatial data collection capabilities previously available only to organizations with substantial financial resources and specialized equipment. Contemporary UAV systems range from consumer-grade platforms to professional mapping-specific configurations, integrating stabilization systems, high-resolution cameras, precision Global Positioning System (GPS) positioning, and automated flight control systems that enable systematic image acquisition with reduced operator expertise requirements and lower operational costs compared to traditional aerial platforms (Liénard et al., 2016).

Equipped with various sensors, UAVs enable acquisition of high-resolution spatial data regarding terrain, land cover, and geographic features (Nesbit & Hugenholtz, 2019). Their low-altitude operation and multi-angle image capture support production of high-resolution maps, even with relatively affordable equipment. UAV-acquired aerial imagery provides orthophotos and detailed surface models while supporting applications such as building reconstruction and local planning in regions with limited road connectivity and infrastructure (Clapuyt et al., 2016).

The integration of photogrammetry and SfM techniques in computer vision has advanced processing capabilities for high-resolution UAV imagery, enabling automated 3D reconstruction pipelines that handle camera calibration, pose estimation, and point cloud generation from unordered image collections without requiring manual intervention or prior geometric constraints (Shahbazi et al., 2015). In computer vision and SfM applications, photogrammetry involves feature detection, feature matching across multiple images, and reconstruction of their relative 3D positions, forming the foundation upon which accurate 3D reconstruction pipelines depend for generating precise models.

Matching multiple views of a scene constitutes a significant component in various computer vision problems, including image retrieval, visual localization, Simultaneous Localization And Mapping (SLAM), and 3D reconstruction (Joshi & Patel, 2020; Ma et al., 2021; Ozyesil et al., 2017; Taketomi et al., 2017). Recent advancements in SfM techniques, combined with classical photogrammetry, have enabled creation of accurate and detailed 3D models, driven by innovation in algorithms and computational architectures for processing large datasets (Chen et al., 2023; Liao et al., 2024).

However, integration of UAV technology with 3D reconstruction workflows introduces specific challenges requiring specialized consideration and optimization. UAV imagery characteristics, including image volumes, diverse viewing geometries, scale variations, environmental influences, and computational processing requirements, create demands on feature extraction and matching algorithms that may not be adequately addressed by evaluation methodologies designed for traditional photogrammetric applications (Xu et al., 2014).

1.2. Problem Statement

Feature detectors, descriptors, and matching algorithms identify and describe corresponding points in images across various domains. However, their results vary depending on image characteristics such as viewpoint, scale, illumination, and compression, presenting specific challenges within the context of UAV imagery (Wang et al., 2023).

A fundamental challenge in UAV-based 3D reconstruction involves objective evaluation and selection of optimal algorithm combinations. Traditional approaches to algorithm evaluation often rely on subjective weighting schemes, researcher-defined importance hierarchies, or single-metric assessments that fail to capture multi-dimensional algorithmic performance in UAV applications. This limitation creates a gap between theoretical algorithm capabilities described in academic literature and practical deployment decisions required in operational environments, where practitioners must balance computational efficiency, reconstruction accuracy, cross-platform compatibility, and resource constraints without evidence-based guidance.

Image-based 3D reconstruction using SfM offers a cost-effective and automated approach for generating three-dimensional models from collections of overlapping images. While SfM produces sparse point clouds and estimates camera poses, which are foundational for 3D modeling, the process proves computationally intensive, particularly when dealing with large datasets typically acquired by UAVs. Efficient feature extraction and matching therefore become essential to making SfM pipelines practical for UAV applications.

SfM proves well-suited for processing unstructured aerial images featuring overlap, enabling estimation of sparse 3D point clouds along with corresponding camera positions and orientations. However, these initial sparse point clouds often prove insufficient for achieving complete and high-quality surface reconstructions required in many downstream applications (Lee & Yoo, 2025; Liu et al., 2024). Multi-View Stereo (MVS) or more broadly, Multi-View Geometry (MVG) algorithms are subsequently applied to increase point cloud density and generate detailed surface models. Accurate identification and robust matching of common regions across overlapping images form the foundation upon which the success of these density improvement methods relies. Feature-based image descriptors have proven effective and are widely adopted for establishing these correspondences (Kholil et al., 2021).

UAVs offer cost-effective, low-altitude image acquisition solutions, with entry-level models now available at affordable prices. Their ease of use, maneuverability in diverse environments, and ability to capture high-resolution imagery from unique perspectives have led to widespread deployment in applications such as surveying, mapping, precision agriculture, and environmental monitoring (Amaral et al., 2020; Ndlovu et al., 2024). Operating typically at altitudes between 30 and 50 meters, UAVs can efficiently capture detailed imagery of target areas. However, battery limitations typically restrict flight durations to 10 to 40 minutes (Mohsan et al., 2022). This limitation necessitates meticulous route planning to maximize coverage and emphasizes the importance of efficiency in data collection and subsequent computationally demanding processing stages required for 3D reconstruction (Arnaldo et al., 2025). Optimizing feature extraction and matching pipelines is therefore essential to address these efficiency concerns and enable the practical and scalable application of UAV-based 3D reconstruction.

The increasing deployment of UAVs across mobile and edge computing environments adds complexity to algorithm selection decisions. Modern UAV applications require on-device processing capabilities, making cross-platform performance evaluation essential for practical deployment decisions.

1.2.1. Challenges in Efficient 3D Reconstruction Specific to UAV Imagery

Extracting accurate 3D models from UAV imagery presents several challenges that compound the problem of objective algorithm evaluation. The generation of vast amounts of high-resolution images by UAVs often results in massive datasets. This creates substantial challenges, as processing such data with traditional SfM pipelines becomes computationally intensive, time-consuming, and impractical without appropriate algorithm optimization (Luo et al., 2024).

Traditional SfM techniques themselves demand computational resources for their core operations, including feature extraction, matching, and optimization. This demand increases with the volume of data produced by UAVs, leading to extended processing times that make real-time or near-real-time processing impractical without careful algorithm selection (Jiang et al., 2020).

Environmental conditions introduce difficulties that vary across operational scenarios. Varying lighting, occlusions, and atmospheric distortions commonly found in UAV imagery can introduce noise and inaccuracies into feature extraction and matching processes, creating efficiency variations that are difficult to predict without evaluation methodologies (Barba et al., 2019).

Camera calibration variations present another concern. Differences in camera calibration parameters across UAV models and diverse settings can lead to inconsistencies in feature detection and matching, which affects overall reconstruction accuracy. Bundle adjustment plays a central role in alleviating these concerns, but optimal algorithm selection remains important for maintaining reconstruction quality across diverse deployment scenarios (Abeho et al., 2024).

The varied imaging conditions and wide range of object scales within UAV imagery create features of differing characteristics. This variability makes it difficult for traditional extraction methods to perform consistently and poses challenges for feature extraction and matching without objective guidance for algorithm selection based on specific operational requirements.

UAV imagery frequently suffers from noise and artifacts. These issues prove particularly prevalent in low-light or high-contrast environments and can degrade the quality of feature extraction, thereby hindering accurate matching and subsequent reconstruction. The impact of these factors varies significantly across different algorithm combinations, making comprehensive evaluation essential for reliable deployment decisions.

1.2.2. Addressing the Challenges

Reconstructing 3D models from UAV imagery requires balancing accuracy, completeness, and computational speed while ensuring objective, evidence-based algorithm selection. Several key areas require systematic attention through evaluation methodologies.

Development of objective evaluation methodologies proves essential for eliminating subjective bias in algorithm selection decisions. This requires frameworks that can systematically assess algorithm combinations across multiple performance dimensions while maintaining theoretical rigor, empirical validity, and practical relevance for real-world deployment scenarios.

Effective management and processing of large datasets prove necessary. This calls for algorithm evaluation frameworks that can identify optimal combinations for handling vast amounts of data generated by UAVs while ensuring that evaluation methodologies can scale to assess performance across diverse operational scenarios and hardware platforms.

Accurate identification and matching of corresponding points remain essential for establishing geometric relationships and building high-quality 3D models. This requires evaluation frameworks that can reliably identify algorithm combinations that are both robust and computationally efficient in their feature extraction and matching capabilities, supported by statistical validation ensuring consistency across different deployment contexts.

Cross-platform efficiency evaluation becomes important as UAV applications expand into mobile and edge computing environments. Algorithm selection frameworks must account for efficiency variations across different hardware architectures while maintaining consistency in efficiency rankings and providing reliable guidance for deployment decisions across diverse computational constraints.

3D reconstruction pipelines must be adapted to suit the distinct properties of UAV imagery through evidence-based algorithm selection. This includes accommodating characteristics such as oblique viewing angles and variations in scale and resolution, ensuring the pipeline can perform well across diverse UAV capture conditions with objective guidance that eliminates subjective bias in method selection.

Achieving a proper balance among accuracy, completeness, and computational speed proves essential. This necessitates evaluation methodologies that can assess algorithm combinations across multiple performance dimensions, providing flexible approaches supported by statistical validation to meet varied project needs and constraints while maintaining objectivity in selection decisions.

1.3. Objectives

The primary objectives of this research are designed to address the challenges in UAV-based 3D reconstruction through development of objective evaluation methodologies:

Development of CUES Methodology: This research develops a novel objective evaluation framework integrating four distinct unsupervised weighting approaches: entropy weighting based on information theory principles, CRITIC method combining contrast intensity and correlation analysis, PCA weighting emphasizing variance-maximizing dimensional components, and variance-based discrimination power measurement. The arithmetic mean combination of these methodologically diverse approaches creates a robust evaluation framework eliminating subjective bias while maintaining strong theoretical foundations and statistical rigor.

Empirical Evaluation Framework: This research provides systematic examination of 784 combinations of feature detection, description, and matching techniques available in the OpenCV computer vision library (Bradski, 2000). The framework provides insights into strengths and weaknesses of these methods for UAV-based 3D reconstruction, identifying optimal combinations for specific operational scenarios and contributing to evidence-based understanding of their comparative efficiency through statistical validation and cross-dataset consistency assessment.

Cross-Platform Efficiency Analysis and Validation: Computational efficiency of feature extraction methods requires assessment across different hardware platforms, particularly mobile devices and high-performance server configurations. This evaluation identifies scalability patterns and efficiency characteristics across various computational environments, determining feasibility of on-device processing while highlighting efficiency gains achievable with more powerful hardware. Cross-platform ranking consistency is validated using the Oxford and Synthetic benchmark evaluations across Intel Xeon Gold 6130, Snapdragon 855, and Dimensity 7200 Ultra processors to ensure reliability across diverse deployment scenarios.

Statistical Validation and Robustness Assessment: The research implements statistical validation including cross-dataset performance validation. This validation approach ensures that the methodology meets standards of academic rigor while providing practical reliability for real-world deployment decisions, establishing confidence intervals and significance testing for all performance comparisons.

Parameter Optimization and Fair Comparison Protocol: Parameter optimization proves important for ensuring fair and meaningful algorithmic comparisons that reflect true algorithmic capabilities rather than configuration-dependent variations. The research implements systematic parameter optimization protocols that explore parameter space of each evaluated method to determine optimal settings through grid search and validation procedures. This approach reduces impact of suboptimal configurations and ensures that comparisons between different OpenCV algorithms reflect their inherent performance characteristics.

Diverse Dataset Development and Cross-Dataset Validation: Performance evaluation requires development of diverse datasets that capture various operational conditions and enable validation of methodology generalizability. The research creates and utilizes multiple data sources, including a multi-layered AirSim drone dataset simulating realistic flight scenarios with controlled environmental variables, synthetic datasets providing systematic manipulation of specific image characteristics, and established benchmark datasets such as Oxford and Pix4D enabling comparison with existing research findings and cross-dataset consistency validation.

Interactive Visualization and Analysis Frameworks: Interactive visualization and analysis frameworks improve interpretability of complex performance data through Plotly-based diagrams and statistical analysis tools. These interactive, web-browser accessible visualizations present performance metrics and 3D reconstruction results in an intuitive manner, enabling researchers and practitioners to gain insights into algorithm behavior, identify performance patterns across multiple evaluation dimensions, and understand relative strengths and weaknesses of different approaches through dynamic, explorable visualizations supporting evidence-based decision making.

Integration of Alternative Methods into COLMAP (Schönberger & Frahm, 2016) Pipeline: The research expands COLMAP's capabilities by systematically integrating alternative OpenCV feature extraction and matching methods beyond the default Scale-Invariant Feature Transform (SIFT) approach. This integration analyzes impact of different feature extraction and matching techniques on 3D model generation quality, exploring opportunities to improve COLMAP's results through algorithmic diversity and providing a framework for future COLMAP extensions with evidence-based method selection guidance.

Novel Efficiency Scoring and Ranking Methodology: Development of an efficiency scoring system represents a key methodological innovation combining computational efficiency, reconstruction quality, and resource utilization metrics into a unified evaluation framework. This scoring system enables direct comparison of algorithm combinations across multiple efficiency dimensions, providing an objective tool for optimal method selection in specific UAV reconstruction scenarios while accounting for the multi-dimensional nature of algorithmic efficiency in real-world applications.

Open-Source Research Contribution and Reproducibility: The commitment to open-source research contribution ensures broad accessibility and reproducibility of findings through publication of code, datasets, evaluation tools, and results in publicly accessible repositories. This approach encourages collaboration within the research community and enables further advancements in UAV-based 3D reconstruction by providing a validated foundation for future investigations and methodological developments while promoting transparency and scientific reproducibility.

1.4. Contributions

This research makes several contributions to the field of efficient 3D reconstruction from UAV imagery, addressing current challenges and extending practices through methodological innovation and empirical validation:

First Evaluation of OpenCV Algorithm Combinations: The research provides the first systematic evaluation of all 784 possible combinations of feature detection, description, and matching methods available in OpenCV for UAV-based 3D reconstruction. This methodical approach reveals strengths and weaknesses of different techniques across various operational contexts, offering direct comparisons that were previously unavailable in the literature and establishing an empirical foundation for evidence-based algorithm selection.

CUES Methodology: Development of a novel theoretically grounded composite efficiency scoring methodology represents a methodological innovation integrating four distinct unsupervised weighting approaches: entropy, CRITIC, PCA, and variance. This framework eliminates subjective bias in algorithm evaluation while maintaining strong statistical foundations and providing objectivity in algorithm assessment for UAV-based 3D reconstruction applications.

Cross-Platform Efficiency Analysis and Benchmarking: Cross-platform efficiency analysis demonstrates how computational resources affect processing times and algorithm rankings across mobile devices and high-performance server configurations. The research establishes benchmarks for scalability assessment using the Oxford and Synthetic datasets across Intel Xeon Gold 6130 CPU @ 2.10GHz, Snapdragon 855 (OnePlus 7), and Dimensity 7200 Ultra (Xiaomi Redmi Note 13 Pro+) processors, providing practical guidance for deployment decisions in resource-constrained environments with focus on real-world UAV applications where edge computing capabilities prove increasingly important.

Standardized Parameter Optimization and Fair Comparison Protocol: Development of a standardized parameter optimization methodology ensures fair comparisons between different OpenCV algorithms through systematic grid search and validation procedures. This contribution addresses a gap in existing literature where algorithmic comparisons often suffer from suboptimal parameter settings, leading to biased conclusions about relative efficiency and providing a replicable framework for future comparative studies.

Novel Efficiency Metrics and Timing Analysis: The research introduces innovative efficiency metrics, including a standardized inlier timing comparison measuring processing time for one thousand inliers, providing practical perspectives on algorithm efficiency that directly relate to real-world applicability in UAV-based reconstruction scenarios. This metric offers more meaningful efficiency assessment than traditional absolute timing measures by normalizing for reconstruction quality outcomes.

Multi-Layered Synthetic and Real-World Dataset Development: The research introduces a multi-layered AirSim drone dataset generated in a controlled simulation environment using a grid-based survey at multiple altitudes and multiple camera orientations. This provides a controlled yet visually realistic testbed for UAV-based 3D reconstruction evaluation. Additionally, synthetic dataset generation methodologies create controlled test scenarios from single images, enabling systematic evaluation of algorithm stability under various conditions such as different lighting situations and geometric transformations.

Enhanced COLMAP Integration and Pipeline Extension: The research expands COLMAP's capabilities by systematically integrating alternative OpenCV feature extraction and matching methods beyond the default SIFT approach. This integration demonstrates potential for improving 3D reconstruction quality through algorithmic diversity and provides a validated framework for future COLMAP extensions, enabling practitioners to implement optimal algorithm combinations within established reconstruction pipelines.

Statistical Validation Framework: A methodology for statistical and visual comparison combines quantitative metrics with interactive visualizations for objective 3D model quality assessment. This dual approach provides both rigorous statistical measures and intuitive understanding of reconstruction quality through statistical analysis and cross-dataset consistency evaluation, supporting evidence-based decision-making in method selection with confidence intervals and significance testing.

Interactive Web-Based Visualization and Analysis Tools: Development of Plotly-based visualization tools enables web-browser exploration of algorithm efficiency data through dynamic, interactive interfaces. These tools make complex efficiency relationships accessible to researchers and practitioners, supporting understanding of algorithmic behavior across different conditions while promoting collaborative analysis and knowledge sharing through web-accessible platforms.

Validated Open-Source Research Framework: The research establishes a complete open-source framework including validated code, datasets, evaluation tools, and reproducible results published in publicly accessible repositories. This contribution promotes scientific reproducibility and enables the research community to build upon the findings while ensuring transparency in methodology and supporting future advancements in UAV-based 3D reconstruction research through collaborative development and validation.

1.5. Thesis Organization

This thesis is structured into six chapters that guide the reader through the research process of developing objective evaluation methodologies for UAV-based 3D reconstruction. The organizational structure ensures logical progression from theoretical foundations through methodological development to empirical validation and practical application:

Chapter 1. Introduction: This foundational chapter establishes research context by introducing the background of UAV-based 3D reconstruction, articulating the problem of objective algorithm evaluation, defining specific research objectives, and outlining the novel contributions of this work. The chapter defines specific challenges associated with UAV-based 3D reconstruction, emphasizes the need for objective evaluation methodologies, and provides a roadmap for the reader by outlining thesis organization and chapter interconnections.

Chapter 2. Literature Review: This chapter provides analysis of existing research on UAV-based photogrammetry, SfM techniques, feature extraction and matching methodologies, and 3D reconstruction pipelines. It examines principles, algorithms, and applications of various techniques while establishing theoretical context, identifying research gaps, and motivating the need for objective evaluation approaches to improve efficiency and accuracy in 3D reconstruction from UAV imagery. The review addresses deep learning approaches while justifying the continued relevance of traditional method evaluation.

Chapter 3. Material and Method: This chapter details the methodology, materials, and procedures used in the research, with emphasis on theoretical foundations of the composite scoring methodology. It describes datasets created and utilized, including synthetic datasets generated with various transformations, the standard Oxford benchmark dataset, drone datasets from Pix4D projects, curated UAV datasets from multiple real-world projects, and the novel multi-layered AirSim dataset generated synthetically using Microsoft AirSim. The chapter explains conceptual foundations of the four unsupervised weighting approaches and implementation details for feature extraction and matching using OpenCV.

Chapter 4. Experimental Design: This chapter presents the detailed experimental setup and validation framework ensuring scientific rigor and practical relevance. It outlines the methodical combinations of 784 methods tested, including coverage of feature detectors, descriptors, and matching algorithms available in OpenCV, with detailed documentation of parameter optimization protocols ensuring fair comparisons. The chapter specifies dataset processing procedures and formally categorizes efficiency metrics used for evaluation.

Chapter 5. Results and Discussion: This chapter presents and analyzes the empirical results of the experiments, with systematic integration of statistical validation and cross-platform analysis. It examines results across all datasets, presented through detailed statistical analysis, plots, interactive visualizations, and heatmaps that reveal performance patterns across different method combinations. The chapter includes correlation analysis to explore statistical relationships between performance metrics and validation of the composite scoring methodology.

Chapter 6. Conclusion: The final chapter synthesizes research findings, reiterates primary contributions with quantitative evidence, and discusses theoretical and practical implications of the novel evaluation methodology. It revisits key objectives and demonstrates how they were addressed through the systematic evaluation framework, highlighting methodological innovation of the composite efficiency scoring system and its validation through multiple statistical approaches. The chapter acknowledges limitations of the current study while outlining promising future research directions.

Appendices: The appendices provide detailed supplementary materials including statistical results, cross-dataset validation analysis, heatmaps and correlation matrices for all evaluated datasets, bar charts showing computational timing results for all method combinations, cross-platform performance comparisons between mobile devices and server configurations, and visualization of the top-performing and least-efficient algorithm combinations for each dataset. Additionally, they contain detailed tables providing performance metrics for optimal method combinations, selected sparse 3D models with CloudCompare distance visualizations demonstrating reconstruction quality, statistical validation results supporting the reliability and robustness of the proposed methodology, and comprehensive feature matching visualizations for top-performing algorithm combinations across all five datasets.

Each chapter builds systematically upon previous foundations while contributing unique insights and methodological innovations that collectively address the central research objectives through empirical validation and statistical rigor. The organizational structure ensures logical progression from theoretical foundations through methodological development to empirical validation and practical application, providing readers with understanding of both the technical innovations and their practical implications for UAV-based 3D reconstruction applications. The thesis maintains consistent academic rigor throughout while balancing theoretical depth with practical relevance, ensuring that contributions advance both scientific understanding and practical capabilities in UAV-based 3D reconstruction applications through the integration of statistical validation, cross-platform efficiency analysis, and objective algorithm evaluation methodologies.

2. LITERATURE REVIEW

This chapter reviews existing literature relevant to UAV-based 3D reconstruction. It examines UAV-based photogrammetry, then explains core components of 3D reconstruction: feature detection and description, feature matching, SfM, and MVS. The chapter reviews existing evaluation methodologies and their limitations, establishing the need for objective evaluation frameworks. Finally, it identifies gaps in algorithm evaluation that this research addresses.

2.1. UAV-based Photogrammetry

UAVs have become important tools for data acquisition and 3D reconstruction across applications. Their maneuverability, cost-effectiveness, and ability to capture high-resolution imagery make them well-suited for tasks where traditional aerial or satellite platforms prove impractical or expensive. This section examines key UAV applications, discusses characteristics of UAV imagery including nadir and oblique perspectives, and reviews the current state of UAV-based datasets, highlighting existing gaps and potential areas for future development.

2.1.1. Overview of UAV Applications

UAV-based photogrammetry has altered data collection and analysis methods in diverse fields. In environmental monitoring and management, UAVs equipped with multispectral or hyperspectral cameras monitor vegetation health, assess forest resources, and map coastal areas (Medvedev et al., 2020). The high-resolution imagery enables detailed analysis of environmental changes, supporting conservation efforts and informed decision-making.

In precision agriculture, UAVs monitor crop health, estimate yields, and optimize resource allocation. High-resolution imagery allows early detection of stress or disease in crops, enabling timely interventions and maximizing productivity (Venkata Sai Chakradhar Reddy et al., 2025). For infrastructure inspection, UAVs offer safer and more efficient alternatives to manual inspections of bridges, power lines, and pipelines. Detailed imagery captured by UAVs identifies structural defects, assesses damage, and plans maintenance activities.

In archaeology, UAVs have changed documentation and site monitoring, allowing creation of accurate 3D models of historical sites and structures, helping with preservation and providing researchers with valuable analysis tools (Mendu & Mbuli, 2025; Partama, 2025). In urban planning and 3D city modeling, UAVs capture high-resolution imagery of buildings, streets, and vegetation, supporting development of accurate virtual representations of urban environments (Muhmad Kamarulzaman et al., 2023).

Recent technological advancements have improved UAV capabilities for 3D reconstruction applications. Modern UAVs integrate sensor fusion systems, combining RGB cameras with LiDAR, thermal, and multispectral sensors for data acquisition (Ulvi & Hamal, 2025). RTK-GPS systems in consumer-grade UAVs have improved georeferencing accuracy, enabling centimeter-level positioning that improves reconstruction quality (Schichler et al., 2025). Battery technology improvements have extended flight times to 45-60 minutes for many commercial UAVs, allowing for larger area coverage and more substantial data collection. Artificial intelligence integration has altered autonomous flight capabilities, with UAVs now capable of intelligent path planning that optimizes overlap patterns for 3D reconstruction, while obstacle avoidance systems enable safer operations in complex environments. The democratization of UAV technology has led to widespread adoption across various industries, with improvements in camera stabilization systems and automated flight modes making high-quality data acquisition accessible to non-expert users, accelerating the development of standardized workflows and best practices.

2.1.2. Characteristics of UAV Imagery

UAV imagery possesses distinct characteristics that influence 3D reconstruction workflows (Nex & Remondino, 2014). UAVs can capture images from various viewpoints and altitudes, including nadir (vertical) and oblique (angled) perspectives. This flexibility allows for scene coverage but introduces challenges in image processing and feature matching due to geometric distortions (Flores-de-Santiago et al., 2020).

To ensure complete scene coverage and accurate 3D reconstruction, UAV image acquisition often involves high overlap between adjacent images. This extensive overlap, including both frontlap (overlap in flight direction) and sidelap (overlap between adjacent flight paths), benefits robust feature matching and accurate geometric correction, although it increases computational demands of processing.

Due to variations in flight altitude and terrain changes, UAV images can exhibit variations in scale and resolution within a single dataset. These variations pose challenges for feature extraction and matching algorithms, which need robustness to such changes (Maes, 2025; Wang et al., 2025). UAV imagery spatial resolution, often expressed as Ground Sampling Distance (GSD), is influenced by flight altitude, with lower altitudes generally leading to higher resolution imagery valuable for capturing detailed features. Accurate calibration of UAV camera intrinsic parameters remains essential for precise 3D reconstruction (Abeho et al., 2024).

UAV imagery can be affected by environmental influences, including varying lighting conditions, weather conditions such as wind and rain, and atmospheric distortions. These factors introduce noise and artifacts into images, impacting overall image quality and subsequent 3D reconstruction processes (Kedzierski & Wierzbicki, 2015; Leem et al., 2025).

2.1.3. Review of UAV-Based Datasets and Gaps

The increasing use of UAV-based photogrammetry has led to development of numerous datasets, providing resources for algorithm development, benchmarking, and research. However, gaps remain in availability and diversity of UAV datasets.

One challenge involves limited availability of large-scale datasets. Many existing datasets focus on specific objects or limited areas. Large-scale datasets covering diverse environments and objects prove essential for evaluating algorithms, developing robust processing pipelines, and addressing computational challenges associated with real-world applications. Creation of synthetic datasets has emerged as a valuable approach to address these limitations (Song, 2025).

Another gap involves the lack of ground truth data. Accurate ground truth data proves important for validating 3D reconstruction algorithm accuracy. However, obtaining high-quality ground truth data for UAV datasets, particularly for large and complex environments, can be expensive and unfeasible. This limits ability to rigorously evaluate algorithm performance and hinders development of accurate and reliable 3D models. Synthetic datasets offer advantages in this regard, allowing automatic and precise generation of ground truth information, including accurate camera poses, depth maps, and 3D point coordinates (Man & Chahl, 2022; Marelli et al., 2020).

UAV datasets often lack diversity in environmental conditions, such as varying lighting, weather, and terrain. Algorithms evaluated on datasets with limited environmental diversity may not generalize well to real-world scenarios. Robust algorithms require datasets that capture a wide range of environmental variations. While real-world datasets naturally contain such variations, controlled manipulation of environmental parameters represents a strength of synthetic data generation. Simulation programs such as Microsoft AirSim and AirGen provide tools for creating virtual environments with controllable lighting conditions, weather patterns, and dynamic elements, enabling systematic evaluation of algorithm robustness under diverse conditions (Shah et al., 2017; Vemprala et al., 2023).

As UAV-based photogrammetry expands into new application domains, specialized datasets that address domain-specific challenges become important. For instance, datasets designed for infrastructure inspection should include images of various types of defects and damage, while datasets for precision agriculture may require multispectral imagery and annotations for specific crop types (Radoglou-Grammatikis et al., 2020; Venkata Sai Chakradhar Reddy et al., 2025). The flexibility of synthetic data generation also extends to the creation of these specialized datasets. Simulation environments can be designed to mimic specific application scenarios, including the introduction of simulated defects on infrastructure or the generation of multispectral imagery for virtual agricultural fields, often with precise annotations that would be laborious or impossible to obtain in the real world (Richter et al., 2016).

Furthermore, cross-platform evaluation has emerged as an important consideration in UAV applications. Modern UAV operations require processing capabilities across diverse hardware architectures, from high-performance server environments to resource-constrained mobile devices and edge computing platforms. However, existing datasets and evaluation frameworks rarely address performance implications of cross-platform deployment, creating a gap in understanding how algorithm performance translates across different computational environments (Shah et al., 2017; Vemprala et al., 2023).

Addressing these dataset gaps is important for further advancing the capabilities of UAV-based 3D reconstruction and enabling its full potential. The development of large-scale, diverse, and well-annotated datasets will support the development of more stable, accurate, and efficient 3D reconstruction algorithms, enabling a wider range of applications and driving innovation in the field. Equally important is the development of evaluation methodologies that can objectively assess algorithm performance across multiple criteria and deployment scenarios.

2.2. Feature Detection and Description

This thesis evaluates feature extraction methods available in OpenCV, including algorithms capable of both detection and description, as well as dedicated detector and descriptor implementations. Table 2.1 systematically lists these methods alongside their respective types and authors.

The computer vision community has increasingly adopted deep learning-based approaches that challenge hand-crafted methods while introducing new considerations for practical deployment. Contemporary research has produced architectures including SuperPoint (DeTone et al., 2018) for self-supervised learning, D2-Net (Dusmanu et al., 2019) combining detection and description in a single network, and LoFTR (Sun et al., 2021) establishing correspondences directly between image pairs using transformer architectures. Despite these advances, evaluation of OpenCV methods remains important for several reasons that directly address practical UAV deployment requirements.

Hand-crafted feature extraction methods require lower computational resources compared to deep learning approaches, operating efficiently on CPU-only systems without GPU acceleration or complex neural network inference frameworks. This computational efficiency proves particularly important for UAV applications where processing must occur on resource-constrained mobile hardware or edge computing platforms. The absence of deep learning framework dependencies, model management requirements, version constraints, and GPU driver compatibility issues translates to improved system reliability and predictable behavior for operational systems where downtime can have consequences.

Established and well-validated 3D reconstruction pipelines such as COLMAP, OpenMVG, and commercial photogrammetry software rely on hand-crafted feature extraction methods as their foundation. These systems have undergone extensive validation across diverse scenarios and represent proven solutions for production environments, while integration of deep learning methods into these pipelines remains an active research area with ongoing challenges. Furthermore, hand-crafted methods demonstrate consistent performance across diverse computational architectures, from high-performance servers to mobile devices and embedded systems, whereas deep learning approaches often exhibit platform-specific variations requiring specialized hardware optimization.

Understanding capabilities and limitations of OpenCV methods provides baseline performance metrics enabling practitioners to make evidence-based decisions about algorithm selection. Comparative studies have demonstrated that while learned features show improvements in discriminative power, hand-crafted descriptors remain competitive in many practical scenarios, particularly when considering computational efficiency and deployment constraints (Schonberger et al., 2017). This evaluation determines when the additional complexity and computational requirements of deep learning approaches are justified versus when hand-crafted methods provide sufficient performance for specific applications.

This research contributes a rigorous, objective evaluation framework that serves multiple purposes: establishing baseline performance metrics enabling meaningful comparison with emerging deep learning approaches, addressing gaps in evaluation methodologies that remain relevant regardless of the underlying feature extraction paradigm, and providing practical guidance for algorithm selection in resource-constrained UAV applications where hand-crafted methods may remain the optimal choice.

Table 2.1 Feature Extraction Methods with Type and Authors

Method	Type	Author
SIFT	Detector & Descriptor	(Lowe, 2004)
ORB	Detector & Descriptor	(Rublee et al., 2011)
BRISK	Detector & Descriptor	(Leutenegger et al., 2011)
KAZE	Detector & Descriptor	(Fernández Alcantarilla et al., 2012)
AKAZE	Detector & Descriptor	(Alcantarilla et al., 2013)
GFTT	Detector	(Jianbo Shi & Tomasi, 1994)
HARRIS LAPLACE	Detector	(Mikolajczyk & Schmid, 2004)
MSER	Detector	(Matas et al., 2004)
FAST	Detector	(Rosten & Drummond, 2006)
STAR	Detector	(Agrawal et al., 2008)
AGAST	Detector	(Mair et al., 2010)
TBMR	Detector	(Xu et al., 2014)
MSD	Detector	(Tombari & Stefano, 2014)
DAISY	Descriptor	(Tola et al., 2010)
BRIEF	Descriptor	(Calonder et al., 2010)
FREAK	Descriptor	(Alahi et al., 2012)
LUCID	Descriptor	(Ziegler et al., 2012)
BOOST	Descriptor	(Trzcinski et al., 2013)
VGG	Descriptor	(Simonyan et al., 2014)
LATCH	Descriptor	(Levi & Hassner, 2016)
BEBLID	Descriptor	(Suárez et al., 2020)
TEBLID	Descriptor	(Suárez et al., 2021)

2.2.1. Feature Detectors

Feature detectors identify distinctive points or regions in images that can be reliably matched across different views. These points, often referred to as keypoints or interest points, are typically defined by high information content, such as corners, blobs, or junctions, which remain stable under various imaging conditions. This section reviews these methods in chronological order, focusing on their detection mechanisms.

Harris Corner Detector (1988):

The Harris Corner Detector represents one of the foundational algorithms in computer vision for identifying distinctive corner points in images (Harris & Stephens, 1988). This method operates on the principle that corners exhibit significant intensity variations in multiple directions, making them important candidates for reliable feature points that can be consistently detected across different viewpoints and imaging conditions.

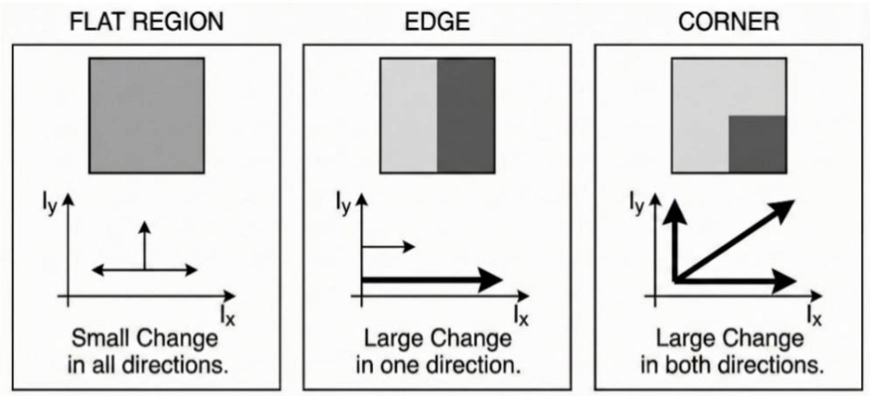


Figure 2.1 Harris Corner Detector with Harris Response

The algorithm constructs a second-moment matrix, also known as the structure tensor, for small windows across the image. This matrix captures spatial derivatives of image intensity in both horizontal and vertical directions, providing a mathematical representation of local intensity variation patterns. The eigenvalues of this matrix reveal the underlying geometric structure: flat regions exhibit small eigenvalues in both directions, edges show one large and one small eigenvalue, while corners demonstrate large eigenvalues in both directions.

The Harris detector employs a computationally efficient corner response function that combines these eigenvalues into a single measure, avoiding explicit eigenvalue computation. This response function incorporates a sensitivity parameter that controls the detector's responsiveness to corner-like structures versus edge-like features. Local maxima above a predetermined threshold are selected as keypoints, ensuring that only the most distinctive corner points are retained.

The detector demonstrates robustness to common image transformations including illumination changes, small rotations, and additive noise, making it suitable for real-world applications where imaging conditions may vary. However, a significant limitation lies in its lack of built-in scale invariance, as the detector operates at a fixed scale determined by the window size used for computing the structure tensor. This limitation makes the basic Harris detector less suitable for applications involving significant scale variations, such as aerial imagery captured at different altitudes.

Despite this limitation, the Harris detector's computational efficiency, conceptual simplicity, and proven reliability have established it as a cornerstone algorithm in computer vision. Its influence extends beyond direct application, as many subsequent corner detection methods have built upon its fundamental principles while addressing its scale invariance limitations. In UAV applications, where computational resources may be limited and real-time performance is important, the Harris detector remains valuable for scenarios involving relatively stable imaging conditions and controlled scale variations.

GFTT (Good Features to Track, 1994):

Good Features to Track (GFTT) represents a refinement of the Harris Corner Detector that emerged from research into feature tracking applications (Jianbo Shi & Tomasi, 1994). The algorithm was specifically designed to identify features that demonstrate stability and trackability across image sequences, making it valuable for applications such as optical flow, object tracking, and motion analysis.

The innovation of GFTT lies in its modified corner quality assessment criterion. While the Harris detector employs a composite measure combining both eigenvalues of the structure tensor, GFTT focuses exclusively on the smaller eigenvalue of the second-moment matrix. This approach is based on the mathematical insight that the smaller eigenvalue represents the minimum rate of intensity change around a pixel, and pixels with large minimum change rates are naturally more distinctive and stable for tracking purposes.

The algorithm's selection criterion states that a pixel qualifies as a "good feature to track" when both eigenvalues of its structure tensor exceed a predefined minimum threshold, with the smaller eigenvalue serving as the primary quality measure. This approach ensures that selected features exhibit strong intensity gradients in all directions, making them highly distinctive and less susceptible to ambiguous matching situations.

GFTT demonstrates improved performance in feature tracking applications compared to the original Harris detector, particularly in terms of tracking accuracy and longevity. The features selected by GFTT tend to maintain their distinctiveness across longer image sequences and under various transformations, making them more reliable for applications requiring temporal consistency. The algorithm's emphasis on the minimum eigenvalue effectively filters out features that might appear distinctive in one direction but lack sufficient contrast in perpendicular directions.

The computational efficiency of GFTT remain comparable to the Harris detector, as it operates on the same fundamental structure tensor computation while applying a different selection criterion. This efficiency makes it suitable for real-time applications where rapid feature detection is essential. However, like its predecessor, GFTT does not naturally address scale invariance, requiring multi-scale extensions or pyramid-based approaches when dealing with significant scale variations.

In UAV-based applications, GFTT's improved tracking capabilities prove particularly valuable for temporal analysis, motion estimation, and change detection scenarios where maintaining feature correspondence across multiple frames is important. The algorithm's reliability in selecting stable features makes it a suitable choice for applications requiring consistent feature detection across varying environmental conditions typical in aerial imaging scenarios.

SIFT Detector (2004):

The Scale-Invariant Feature Transform (SIFT) detector represents an advancement in feature detection that addresses the important limitation of scale invariance (Lowe, 2004). SIFT addresses scale invariance through sophisticated scale-space analysis, establishing itself as one of the most influential feature detection algorithms in computer vision.

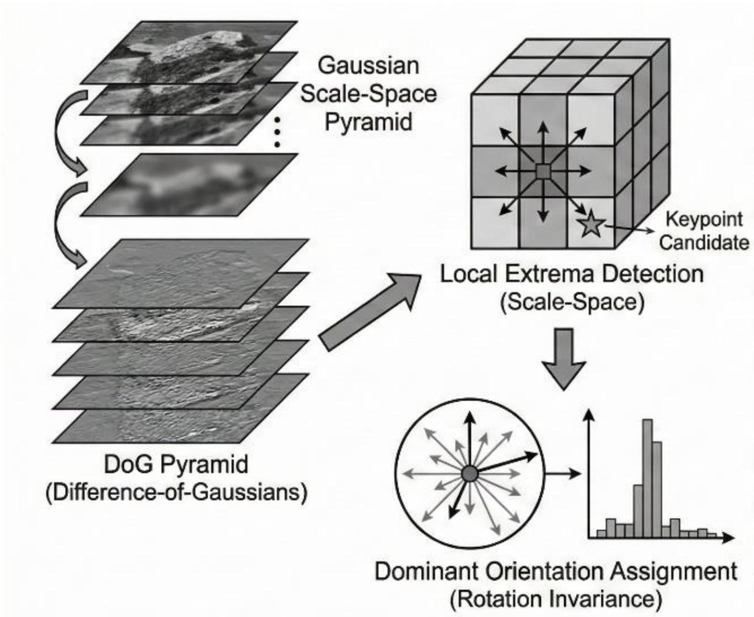


Figure 2.2 SIFT Detector Concept

The algorithm's innovation lies in its scale-space analysis approach, which systematically examines image content across multiple scales to identify features that remain stable regardless of viewing distance or object size. SIFT constructs a scale-space representation by convolving the original image with Gaussian kernels of progressively increasing standard deviations, creating a pyramid of smoothed images that capture structural information at different levels of detail.

SIFT employs a Difference-of-Gaussians (DoG) approximation to the Laplacian-of-Gaussian operator for keypoint identification. This approach involves computing differences between adjacent Gaussian-smoothed images within the scale-space pyramid, creating a DoG pyramid that highlights blob-like structures at various scales. Keypoint candidates are identified as local extrema in this three-dimensional space, considering both spatial location and scale dimension simultaneously.

The algorithm incorporates keypoint refinement procedures that ensure stability and accuracy of detected features. Initial keypoint candidates undergo sub-pixel localization using quadratic interpolation, improving positional accuracy beyond the discrete pixel grid. SIFT applies filtering criteria to eliminate low-contrast points and edge responses that might compromise feature distinctiveness.

A crucial component of SIFT's stability is its orientation assignment mechanism, which provides rotation invariance by identifying dominant gradient orientations within the keypoint neighborhood. This process analyzes gradient magnitude and direction in a circular region around each keypoint, constructing orientation histograms that reveal predominant directional patterns. The peak orientation becomes the keypoint's canonical direction, ensuring consistent feature detection regardless of image rotation.

SIFT's multi-scale detection capability makes it suitable for UAV applications, where imaging conditions frequently involve significant scale variations due to altitude changes, zoom adjustments, or varying object distances. The algorithm's proven robustness to rotation, scale, and illumination changes has established it as a gold standard for feature detection in challenging real-world scenarios typical of aerial imaging applications.

Harris-Laplace Detector (2004):

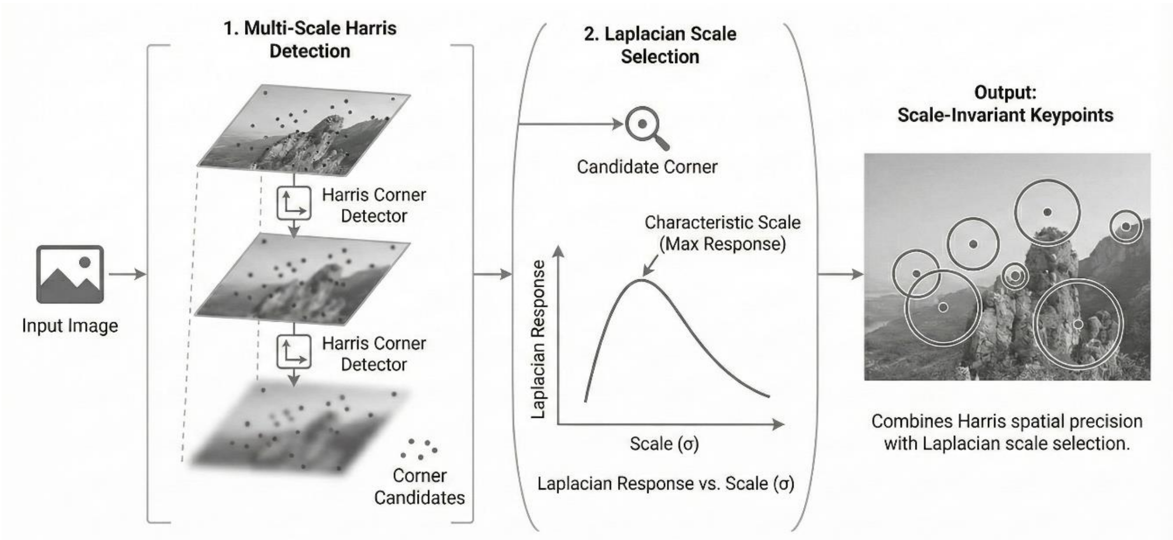


Figure 2.3 Harris-Laplace Detector Concept

The Harris-Laplace detector represents a hybrid approach that combines the precise localization capabilities of the Harris corner detector with scale selection mechanisms to achieve scale-invariant feature detection (Mikolajczyk & Schmid, 2004). This algorithm emerged from the recognition that while the Harris corner detector provides spatial accuracy and reliability for corner detection, its limitation regarding scale invariance restricts its applicability in real-world scenarios involving varying viewing distances or imaging scales.

The algorithm's dual-stage detection process addresses these limitations through integration of two mathematical frameworks. The first stage employs Harris corner detection across multiple scales within a Gaussian scale-space pyramid, ensuring that corner features can be detected regardless of their characteristic scale. This multi-scale Harris detection provides spatial coverage while maintaining the precision and reliability of Harris corners.

The second stage implements scale selection through Laplacian-based analysis that identifies the most characteristic scale for each detected corner. This process involves computing Laplacian of Gaussian operator responses at various scales centered on each corner location and searching for local maxima in the scale dimension. The scale that produces the maximum Laplacian response indicates the most stable and representative scale for that particular corner feature.

Laplacian-based scale selection provides advantages over alternative scale determination methods. The Laplacian operator's sensitivity to blob-like structures and scale-dependent characteristics enables accurate identification of the scale at which corner features exhibit maximum structural coherence. This mathematical foundation ensures that selected scales correspond to genuine structural characteristics rather than artifacts of particular scale choices.

Spatial localization accuracy represents a fundamental strength inherited from the underlying Harris corner detector, ensuring that detected features maintain precise spatial coordinates despite the additional scale analysis. The algorithm preserves the sub-pixel localization capabilities that characterize high-quality corner detection while extending these capabilities into the scale dimension.

In UAV applications, Harris-Laplace proves particularly valuable for scenarios involving altitude variations, zoom changes, or applications requiring consistent feature detection across varying imaging scales. The algorithm's combination of spatial precision and scale invariance makes it suitable for precision navigation, structure-from-motion applications, and detailed mapping tasks where reliable corner features are essential for accurate geometric computations.

MSER Detector (2004):

Maximally Stable Extremal Regions (MSER) represents a different approach to feature detection that identifies stable blob-like regions rather than point features (Matas et al., 2004). This method emerged from the recognition that many visual elements in real-world scenes, such as text characters, logos, and distinctive objects, exhibit coherent regions with consistent properties that can be more reliably detected and tracked than isolated corner points.

The algorithm's innovation lies in its region-based analysis approach, which systematically examines connected components across a range of intensity thresholds. MSER begins by considering the image as a topographic surface where pixel intensities represent elevation values. By progressively varying a global intensity threshold, the method tracks how connected regions grow and shrink, identifying areas that demonstrate stability across significant threshold ranges.

The concept of "maximal stability" refers to regions whose area changes minimally as the intensity threshold varies within a substantial range. This stability criterion ensures that detected regions represent genuine structural features rather than artifacts of particular threshold choices. The algorithm quantifies stability by measuring the rate of area change relative to threshold variation, selecting regions where this rate reaches local minima.

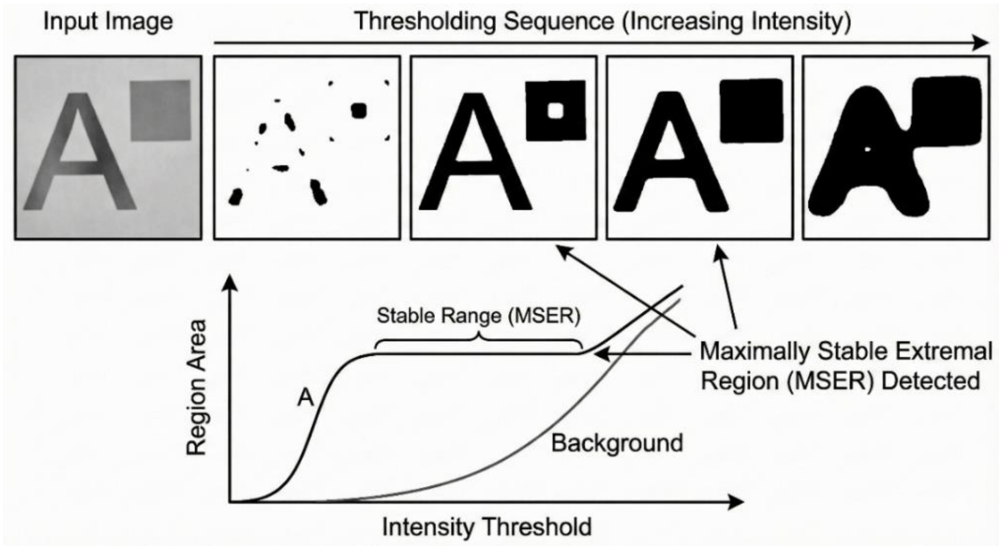


Figure 2.4 MSER Detector Concept

MSER's region-based approach provides advantages over point-based detection methods. The detected regions naturally provide shape information beyond simple location coordinates, supporting more sophisticated matching and recognition algorithms. These regions demonstrate stability under various geometric and photometric transformations, including significant viewpoint changes, illumination variations, and partial occlusions.

The algorithm exhibits stability to affine transformations, making it valuable for applications involving perspective distortions common in aerial imagery. MSER regions maintain their essential characteristics under stretching, rotation, and skewing transformations, providing stable features for geometric verification and camera pose estimation.

MSER demonstrates good performance in detecting text, symbols, and structured elements within images, making it useful for applications involving man-made environments or document analysis. However, MSER can be sensitive to image noise and compression artifacts, which may create spurious connected components. The method may also struggle in highly textured natural environments where stable regions are less prevalent.

In UAV applications, MSER proves particularly valuable for scenarios involving urban environments, infrastructure inspection, or archaeological documentation where structured elements and man-made objects provide natural region-based features. The algorithm's ability to detect distinctive shapes and symbols makes it useful for landmark recognition and navigation applications in GPS-denied environments.

FAST Detector (2006):

Features from Accelerated Segment Test (FAST) improved real-time feature detection by introducing a simple yet effective corner detection algorithm specifically designed for applications requiring computational efficiency (Rosten & Drummond, 2006). FAST emerged from the recognition that many computer vision applications, particularly those involving real-time processing or resource-constrained environments, needed feature detection capabilities that prioritized speed without completely sacrificing detection quality.

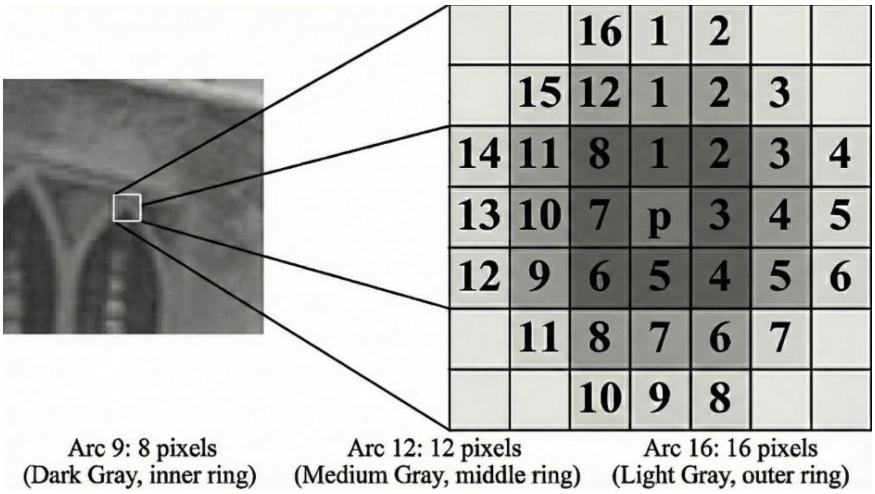


Figure 2.5 FAST detector arcs 9, 12, 16

The algorithm's innovation lies in its streamlined corner detection mechanism that eliminates the complex matrix computations required by traditional methods such as Harris or SIFT. Instead, FAST employs a direct intensity comparison approach that evaluates the relationship between a central pixel and its surrounding neighborhood. The method examines a circular pattern of 16 pixels positioned around the candidate point, creating a test that can rapidly determine corner likelihood.

The core decision criterion requires that a contiguous arc of n pixels (typically 9, 12, or 16) around the circle must all be either significantly brighter or significantly darker than the central pixel by a predetermined threshold value. This simple binary test captures the essential characteristic of corner points - the presence of rapid intensity transitions in multiple directions - while avoiding computationally expensive gradient calculations or eigenvalue decompositions.

FAST's computational efficiency stems from several algorithmic optimizations. The method can employ early termination strategies by testing only a subset of the circular pixels initially, allowing rapid rejection of obviously non-corner pixels. Additionally, the algorithm can be implemented using bit-wise operations and lookup tables, making it suitable for embedded systems and mobile computing platforms where processing power and energy consumption are important constraints.

The detector demonstrates considerable speed advantages over traditional methods, often achieving processing rates significantly faster than gradient-based approaches. This speed advantage makes FAST particularly valuable for applications such as visual SLAM, real-time tracking, and autonomous navigation systems where rapid feature detection directly impacts system responsiveness and overall performance.

However, FAST's emphasis on computational efficiency comes with certain limitations. The basic algorithm lacks built-in scale and rotation invariance, making it sensitive to viewpoint changes and scale variations that commonly occur in UAV imagery. Additionally, FAST can be more susceptible to noise and tends to cluster multiple detections around strong corners, requiring post-processing steps such as non-maximum suppression to ensure feature distribution quality.

Despite these limitations, FAST's speed and reasonable detection quality have made it a cornerstone algorithm for many real-time computer vision systems. Its influence extends beyond direct application, as numerous subsequent algorithms have incorporated FAST-based detection as a foundational component while addressing its invariance limitations through multi-scale frameworks or orientation estimation techniques.

STAR (CenSurE) Detector (2008):

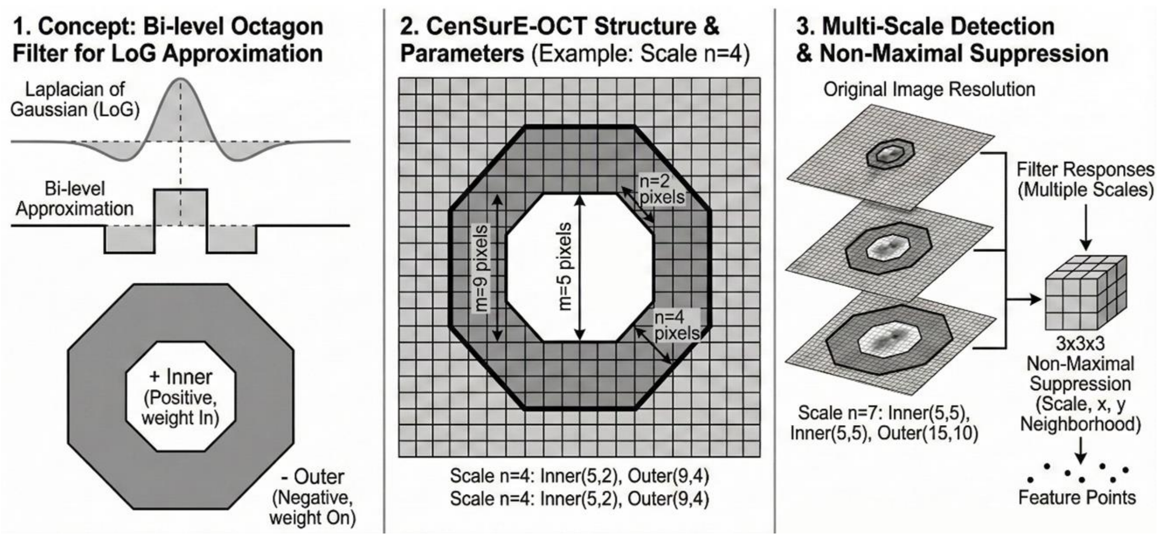


Figure 2.6 CenSurE Detector Bi-level Octagon Filters and Multi-scale Feature Detection

The STAR detector, also known as CenSurE (Center-Surround Extremas), represents an approach to feature detection that prioritizes computational efficiency while maintaining the robustness characteristics essential for reliable feature matching (Agrawal et al., 2008). This detector emerged from analysis of the computational bottlenecks in traditional scale-space methods and implements strategic approximations that reduce processing requirements without significantly compromising detection quality.

The innovation underlying STAR detection lies in its approximation of the Laplacian of Gaussian (LoG) operator using bi-level box filters arranged in distinctive star-shaped spatial patterns. Traditional LoG computation requires expensive convolution operations with Gaussian kernels at multiple scales, creating computational overhead that limits real-time applicability. STAR addresses this limitation by designing box filter approximations that capture the essential characteristics of LoG responses while enabling efficient computation through integral image techniques.

The star-shaped filter configuration represents a balanced approach between computational efficiency and detection accuracy. The filters employ symmetric patterns that approximate the center-surround structure of LoG operators, using positive and negative regions arranged to capture local intensity variations indicative of corner and blob features. This geometric arrangement enables effective feature detection while requiring only simple addition and subtraction operations on integral image values.

Multi-scale detection in STAR employs pyramid-based processing where the star-shaped filters are applied at multiple scales to ensure scale invariance. Rather than constructing explicit scale pyramids through image resampling, STAR achieves multi-scale analysis by varying the size of the star-shaped filters while operating on the original image resolution. This approach reduces memory requirements and computational overhead compared to traditional pyramid-based methods.

Performance characteristics demonstrate that STAR achieves good computational efficiency while maintaining competitive detection quality compared to more sophisticated alternatives. The algorithm's speed advantages make it particularly suitable for applications requiring real-time processing or handling large image sequences where computational efficiency is important.

In UAV applications, STAR's computational efficiency makes it valuable for onboard processing scenarios where computational resources are limited and real-time performance is important. The algorithm's speed enables continuous processing of video streams from UAV cameras, supporting applications such as real-time navigation, obstacle avoidance, and autonomous flight control where immediate feature detection is essential for safe operation.

AGAST Detector (2010):

Adaptive and Generic Accelerated Segment Test (AGAST) represents an evolutionary advancement of the FAST corner detection algorithm that addresses its fundamental limitations while preserving its computational efficiency characteristics (Mair et al., 2010). AGAST emerged from detailed analysis of FAST's performance across diverse imaging conditions and systematic optimization of its decision-making processes to achieve improved robustness without sacrificing the speed advantages that made FAST attractive for real-time applications.

The core innovation underlying AGAST lies in its adaptive decision tree architecture that optimizes the sequence of pixel intensity comparisons based on local image characteristics and learned patterns from extensive training datasets. Unlike FAST's fixed comparison pattern, AGAST employs intelligent decision trees that adapt the testing sequence to maximize information gain at each step, enabling more efficient rejection of non-corner pixels while improving detection accuracy for genuine corner features.

The algorithm's adaptive thresholding mechanism represents an advancement over FAST's static threshold approach, enabling stable detection across varying illumination conditions and image qualities. AGAST analyzes local intensity distributions to determine appropriate threshold values that maintain consistent detection sensitivity while adapting to local contrast characteristics. This adaptive approach prevents over-detection in high-contrast regions and under-detection in low-contrast areas that commonly affect fixed-threshold methods.

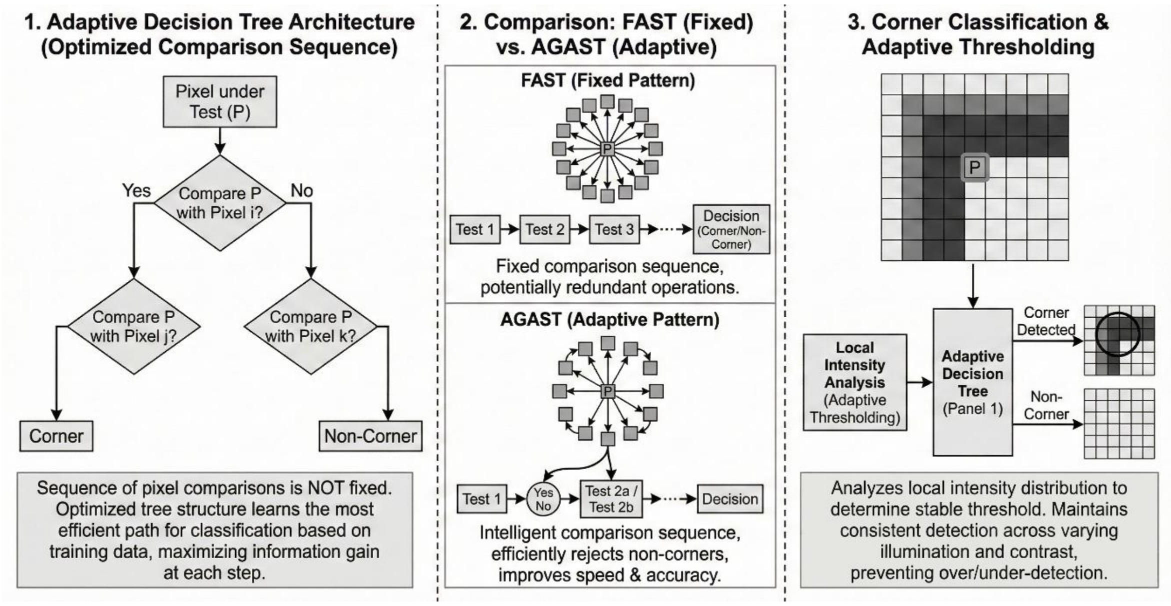


Figure 2.7 AGAST Detector Concept

Decision tree optimization in AGAST employs machine learning techniques to determine optimal pixel comparison sequences that minimize the expected number of operations required for corner classification. The algorithm analyzes large training datasets to identify comparison patterns that provide maximum discriminative power with minimal computational overhead, resulting in decision trees that distinguish corner pixels from edge pixels and homogeneous regions.

The algorithm demonstrates improved repeatability compared to FAST, providing more consistent feature detection across different imaging conditions and geometric transformations. This improved consistency improves matching reliability in applications requiring stable correspondence establishment across image sequences or different viewpoints.

Computational efficiency remains a primary strength of AGAST, preserving the fundamental speed advantages that make corner-based detection suitable for real-time applications. The optimized decision trees and adaptive thresholds maintain low computational overhead while providing improved detection quality.

In UAV applications, AGAST's combination of speed and improved robustness makes it particularly valuable for scenarios involving varying lighting conditions, altitude changes, or applications requiring continuous processing of high-frame-rate video streams. The adaptive thresholding capabilities help maintain consistent detection across the diverse imaging conditions commonly encountered in aerial photography, while the computational efficiency supports onboard processing for autonomous navigation and real-time decision-making systems.

BRISK Detector (2011):

Binary Robust Invariant Scalable Keypoints (BRISK) represents a comprehensive approach to feature detection specifically designed for real-time applications and resource-constrained environments where computational efficiency must be balanced with robust performance (Leutenegger et al., 2011). BRISK emerged from the recognition that many practical computer vision applications, particularly those involving mobile platforms or embedded systems, required feature detection capabilities that could operate effectively within strict computational and power consumption constraints.

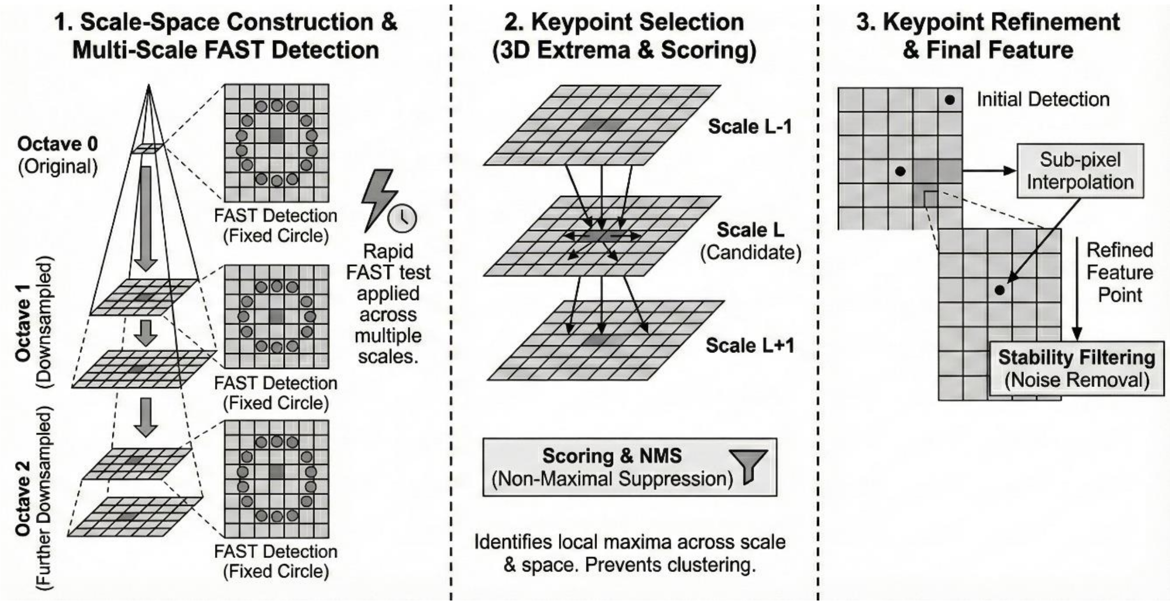


Figure 2.8 BRISK Detector Concept

The algorithm's detection component builds upon the proven efficiency of the FAST corner detection algorithm while systematically addressing its critical limitations regarding scale invariance. BRISK employs a multi-scale FAST detection strategy, applying the rapid corner detection test across several octaves of an image pyramid to capture features at different scales. This pyramid-based approach ensures that features can be detected consistently regardless of viewing distance or object size variations.

The scale-space construction process involves creating multiple octaves of the original image, each representing a different scale level. Within each octave, BRISK applies the FAST corner detection algorithm to identify potential keypoints based on the rapid intensity comparison tests around circular neighborhoods. This multi-scale application ensures comprehensive coverage of potential feature scales while maintaining the computational efficiency advantages of the underlying FAST algorithm.

Keypoint selection employs a sophisticated scoring mechanism that identifies local maxima in the FAST corner response across both spatial and scale dimensions. This three-dimensional extrema detection ensures that selected keypoints represent genuine scale-space features rather than artifacts of particular scale levels. The scoring process incorporates non-maximum suppression techniques to prevent clustering of multiple detections around the same physical feature.

BRISK incorporates advanced keypoint refinement procedures that improve localization accuracy and stability. The algorithm employs sub-pixel interpolation techniques to enhance positional precision beyond the discrete pixel grid, while stability filtering removes keypoints that may be sensitive to noise or poorly localized. These refinement steps ensure that the final feature set maintains high quality despite the rapid detection process.

The detector's scale invariance mechanism operates through careful analysis of the FAST response across different scale levels. Features that demonstrate consistent responses across multiple scales are prioritized, ensuring that detected keypoints correspond to stable structural elements rather than scale-dependent artifacts. This multi-scale consistency analysis provides robust scale invariance while maintaining computational efficiency.

BRISK's design philosophy emphasizes practical deployment considerations, incorporating features that enhance robustness in real-world scenarios. The algorithm demonstrates good performance under moderate geometric transformations, illumination changes, and image compression effects commonly encountered in practical applications. The balance between speed and robustness makes BRISK particularly suitable for applications requiring consistent real-time performance.

In UAV applications, BRISK's combination of speed and scale invariance makes it valuable for scenarios involving altitude variations, zoom changes, or applications requiring rapid processing of large image sequences. The algorithm's efficiency proves particularly beneficial for onboard processing scenarios where computational resources and power consumption are critical constraints, while its scale invariance handles the varying perspectives common in aerial imaging applications.

ORB Detector (2011):

Oriented FAST and Rotated BRIEF (ORB) represents an achievement in creating a completely patent-free, high-performance alternative to proprietary feature detection and description algorithms (Rublee et al., 2011). ORB emerged as a response to the patent restrictions surrounding SIFT and SURF, providing the computer vision community with an open-source solution that combines computational efficiency with stable performance characteristics suitable for a wide range of practical applications.

The detection component of ORB builds upon the proven efficiency of the FAST detector while addressing its limitation regarding rotation invariance. ORB begins by applying the FAST corner detection algorithm to identify potential keypoints based on the rapid intensity comparison test around circular neighborhoods. This foundation ensures that ORB maintains the computational efficiency that makes FAST suitable for real-time applications and resource-constrained environments.

The algorithm's key innovation lies in its orientation assignment mechanism that transforms FAST's rotation-sensitive detections into rotation-invariant features. ORB employs the intensity centroid method, which calculates the weighted center of mass of pixel intensities within a circular patch around each detected keypoint. This approach computes the centroid by summing intensity-weighted pixel coordinates, creating a vector from the keypoint center to the intensity centroid that indicates the predominant intensity distribution direction.

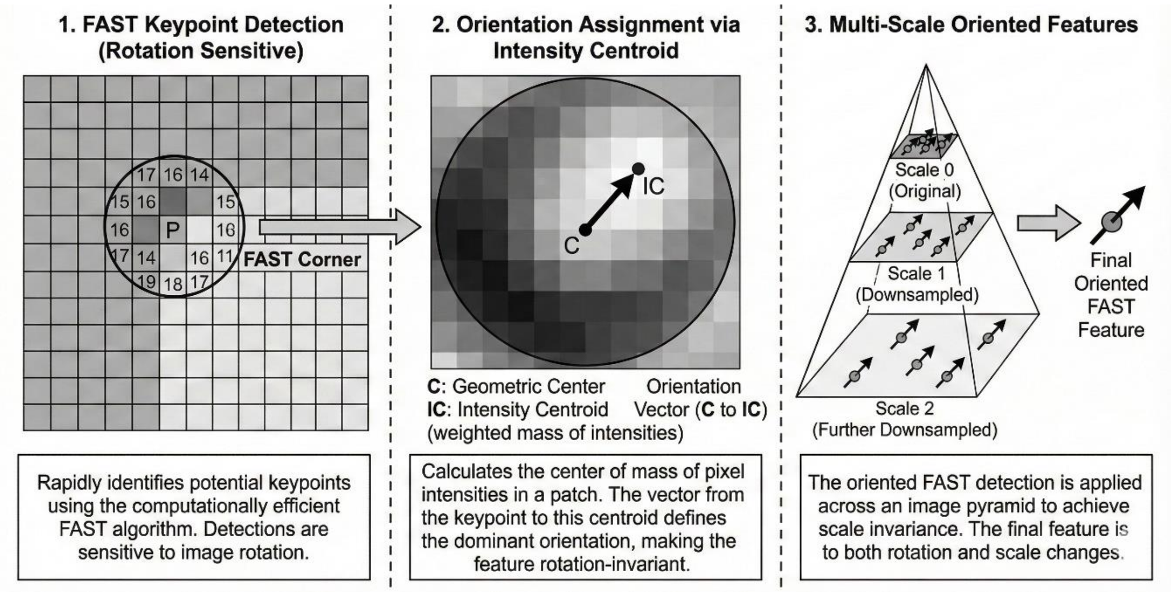


Figure 2.9 ORB Detector Concept

The intensity centroid method provides several advantages over alternative orientation assignment approaches. It operates on the same pixel data used for keypoint detection, avoiding additional computational overhead from gradient calculations or orientation histograms. The method demonstrates good stability under moderate noise conditions and provides consistent orientation estimates across different imaging conditions, making it suitable for practical deployment scenarios.

ORB incorporates multi-scale detection capabilities by constructing an image pyramid and applying the oriented FAST detector at each scale level. This multi-scale approach addresses the scale invariance limitation of the basic FAST detector, enabling ORB to detect features consistently across different viewing distances and object sizes.

ORB's combination of efficiency, patent-freedom, and robust performance has made it one of the most widely adopted feature detection algorithms in practical computer vision systems. Its particular strength in UAV applications stems from its ability to provide reliable feature detection with minimal computational overhead, making it suitable for onboard processing scenarios where power consumption and processing capabilities are limited. The algorithm's rotation invariance proves especially valuable for aerial imaging applications where camera orientation frequently varies due to platform movement or intentional gimbal adjustments.

KAZE Detector (2012):

KAZE represents a paradigm shift in feature detection by introducing nonlinear scale-space analysis that challenges the traditional Gaussian-based approaches (Fernández Alcantarilla et al., 2012). The algorithm emerged from the recognition that conventional scale-space methods, while mathematically elegant, suffer from built-in limitations when processing images containing complex structures with varying levels of detail and noise.

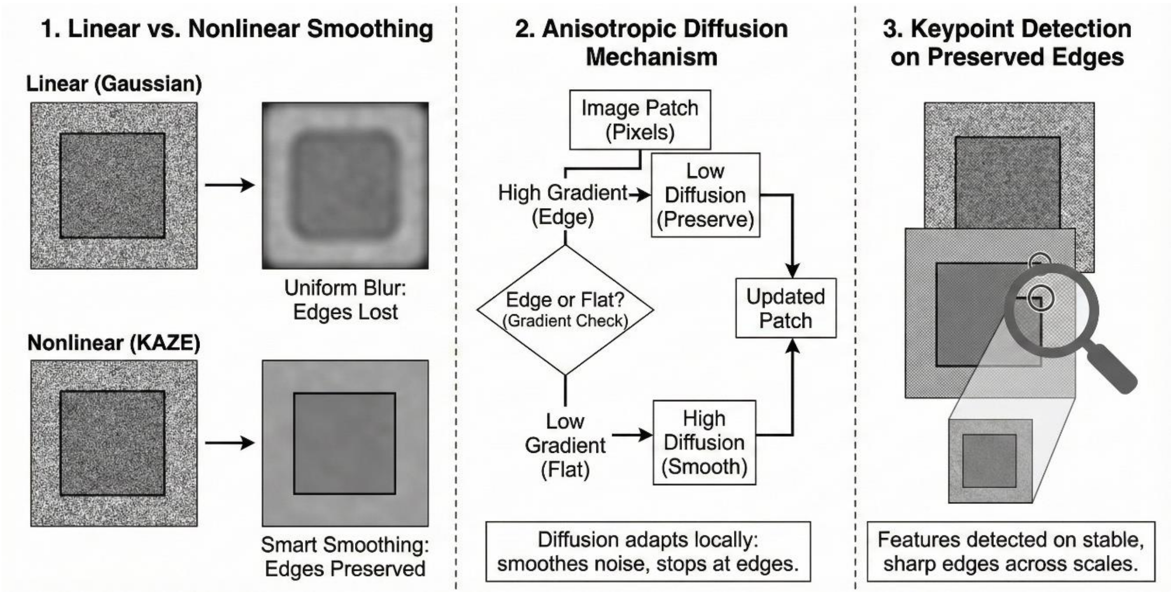


Figure 2.10 KAZE Detector Concept

The algorithm's innovation lies in its construction of nonlinear scale-space representations using anisotropic diffusion filtering instead of traditional Gaussian convolution. While conventional methods apply uniform smoothing across entire images, KAZE employs adaptive diffusion processes that preserve important structural boundaries while smoothing regions of lower information content. This selective processing approach enables the retention of fine details and sharp edges that would be lost in traditional linear scale-space construction.

Anisotropic diffusion filtering operates by solving partial differential equations that govern the flow of information across the image domain. The diffusion process adapts locally to image content, reducing smoothing in regions containing significant gradients while allowing more aggressive noise reduction in homogeneous areas. This content-aware processing creates scale-space representations that maintain structural integrity across multiple scales, providing more reliable foundations for feature detection.

The keypoint detection process identifies extrema within this nonlinear scale-space pyramid, ensuring that detected features correspond to genuinely stable structures rather than artifacts of excessive smoothing. The preserved edge information enables KAZE to detect features in challenging environments where traditional methods might fail, such as images containing fine textures, weak contrasts, or complex geometric patterns.

KAZE's nonlinear approach demonstrates improved performance in scenarios involving image blur, noise, and varying illumination conditions compared to traditional linear methods. The computational complexity of KAZE is higher than traditional methods due to the iterative nature of the diffusion equations, but this computational investment yields improved feature quality and stability in challenging imaging conditions.

In UAV applications, KAZE's enhanced robustness to blur and noise proves particularly valuable for scenarios involving platform movement, atmospheric disturbances, or varying image quality conditions. The algorithm's ability to preserve fine structural details makes it useful for applications requiring precise feature localization, such as infrastructure inspection or archaeological documentation where subtle features may carry significant information.

AKAZE Detector (2013):

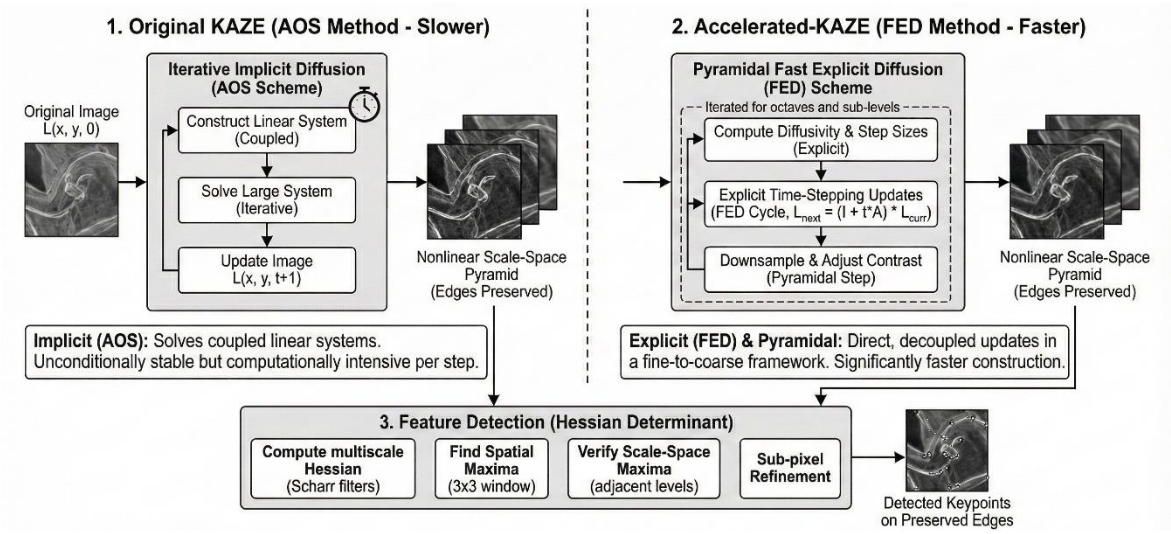


Figure 2.11 AKAZE Detector Concept

Accelerated-KAZE (AKAZE) represents a breakthrough in computational optimization for nonlinear scale-space feature detection that preserves the fundamental advantages of the KAZE algorithm while achieving dramatic improvements in processing speed through innovative mathematical and algorithmic improvements (Alcantarilla et al., 2013). AKAZE emerged from systematic analysis of the computational bottlenecks in nonlinear diffusion processing and implements strategic optimizations that make nonlinear scale-space methods practical for real-time applications and resource-constrained environments.

The advancement underlying AKAZE lies in its adoption of Fast Explicit Diffusion (FED) schemes as a replacement for the computationally intensive Additive Operator Splitting (AOS) methods employed in the original KAZE algorithm. FED schemes solve the same nonlinear diffusion equations that provide KAZE's improved edge-preserving characteristics but employ explicit time-stepping procedures that can be implemented more practically on modern processor architectures while maintaining mathematical accuracy and stability.

Fast Explicit Diffusion implementation supports parallel processing and vectorization optimizations that utilize modern computational hardware capabilities. Unlike implicit methods that require solving systems of linear equations at each time step, FED schemes employ explicit update formulas that can be computed using SIMD instructions and parallel processing frameworks. This computational approach reduces processing time while preserving the essential mathematical properties of nonlinear diffusion.

The nonlinear scale-space construction in AKAZE maintains the same fundamental mathematical framework as KAZE, ensuring that the improved computational efficiency does not compromise the improved structural preservation characteristics that distinguish nonlinear methods from traditional linear approaches. The algorithm constructs anisotropic diffusion pyramids that preserve edge information while smoothing noise.

Performance benchmarking demonstrates that AKAZE achieves processing speeds approaching those of conventional linear-filtering methods while preserving the enhanced robustness and structural preservation capabilities of nonlinear scale-space processing. This breakthrough makes nonlinear methods practical for applications previously limited to simpler algorithms due to computational constraints.

The algorithm maintains good scale and rotation invariance through orientation assignment mechanisms that employ the improved edge preservation provided by nonlinear diffusion. The improved structural representation often enables more accurate orientation estimation compared to traditional linear methods.

In UAV applications, AKAZE provides an optimal balance between computational efficiency and detection quality, enabling robust feature detection for onboard processing scenarios where both accuracy and speed are important. The algorithm's improved performance under challenging imaging conditions makes it particularly valuable for applications involving atmospheric disturbances, motion blur from platform movement, or varying environmental conditions commonly encountered in aerial operations.

TBMR Detector (2014):

Tree-Based Morse Regions (TBMR) represents an approach to feature detection that employs mathematical concepts from Morse theory and computational topology to identify stable image regions based on their topological characteristics rather than conventional gradient-based criteria (Xu et al., 2014). This detector emerged from the recognition that traditional feature detection methods may fail to capture important structural information that can be revealed through topological analysis of image intensity landscapes.

The mathematical framework underlying TBMR employs Morse theory, which analyzes the topological structure of smooth manifolds through critical point analysis. In the context of image processing, the image intensity function defines a two-dimensional manifold where local intensity variations create topological features that can be systematically analyzed. TBMR employs these mathematical principles to identify regions that exhibit stable topological characteristics across different intensity threshold levels.

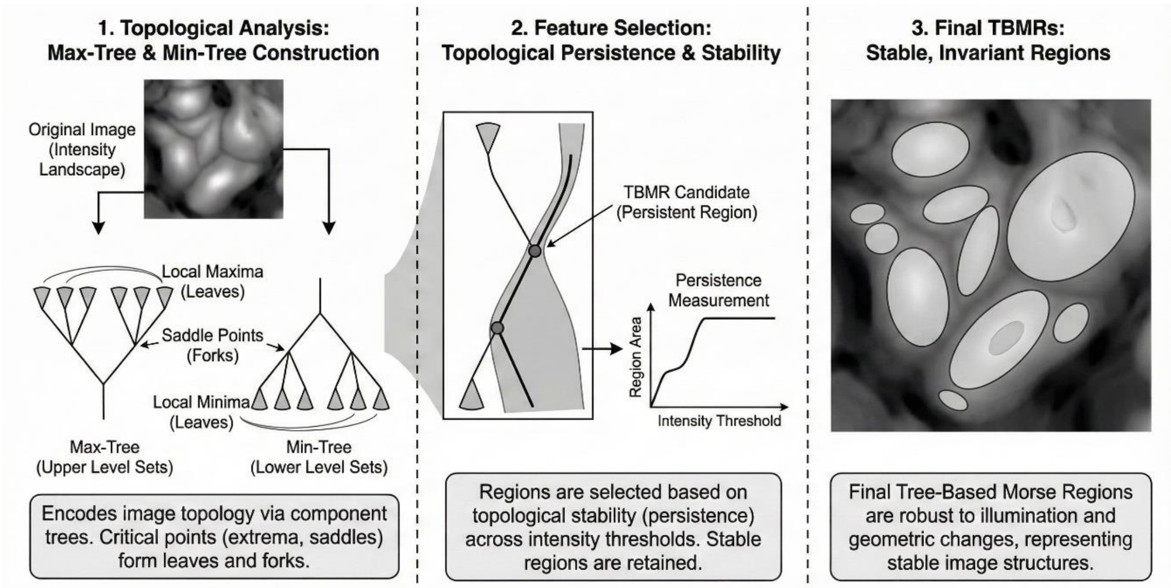


Figure 2.12 TBMR Detector Concept

Tree structure construction forms the core of TBMR's detection methodology, involving the systematic creation of Max-Trees and Min-Trees that encode the hierarchical organization of image intensity levels. These tree structures provide representations of the image's topological organization that remain stable under various imaging transformations.

The persistence measurement methodology quantifies the stability of topological features by analyzing how tree node structures evolve as intensity thresholds change. Regions that maintain their topological characteristics across wide threshold ranges receive high persistence scores, indicating stable image structures. This approach provides objective criteria for feature selection that are stable to noise and minor intensity variations.

In UAV applications, TBMR's robustness to illumination changes and geometric transformations makes it valuable for scenarios involving varying environmental conditions, sun angle changes, and platform orientation variations. The algorithm's ability to detect both high-contrast and subtle features proves useful for applications requiring detailed scene analysis, such as terrain mapping and infrastructure monitoring.

MSD Detector (2014):

Maximal Self-Dissimilarity (MSD) represents a paradigm in feature detection that redefines feature distinctiveness by focusing on local uniqueness characteristics rather than traditional gradient-based criteria (Tombari & Stefano, 2014). This detector emerged from the recognition that the most valuable features for matching are those that exhibit maximum dissimilarity from their surrounding neighborhoods, creating distinctive signatures for reliable identification.

The core principle underlying MSD detection involves the systematic computation of self-dissimilarity measures that quantify how different each image location is from its surrounding neighborhood. Unlike conventional detectors that identify specific structural patterns such as corners or blobs, MSD adopts a general approach that seeks regions demonstrating maximum local uniqueness regardless of their specific geometric characteristics.

Self-dissimilarity computation employs patch comparison algorithms that analyze the statistical and structural differences between local image patches and their surrounding regions. The algorithm defines circular or rectangular neighborhoods around each pixel location and systematically compares the central patch against multiple surrounding patches within a defined search radius.

Multi-scale analysis in MSD ensures coverage of features at different spatial scales by computing self-dissimilarity measures across multiple patch sizes and search radii. This multi-scale approach provides built-in scale invariance without requiring explicit scale normalization procedures.

The uniqueness-based detection criterion provides natural robustness to various imaging transformations, as patches that are genuinely unique within their neighborhoods tend to maintain their distinctive characteristics under moderate geometric and photometric transformations. This robustness emerges naturally from the dissimilarity computation process rather than requiring explicit invariance mechanisms.

The algorithm demonstrates excellent performance in scenarios involving complex textures, irregular patterns, and challenging feature types that may not be effectively captured by conventional geometric feature detectors. The uniqueness-based criterion proves particularly valuable for detecting distinctive regions in natural images, textured surfaces, and complex scenes.

In UAV applications, MSD's emphasis on local uniqueness makes it particularly valuable for scenarios involving natural terrain features, vegetation patterns, or complex urban environments where traditional geometric features may be sparse or unreliable. The algorithm's robustness to illumination changes and moderate geometric transformations proves beneficial for aerial imagery applications involving varying sun angles, atmospheric conditions, and platform perspectives commonly encountered in UAV missions.

2.2.2. Feature Descriptors

Feature descriptors characterize the local appearance of keypoints, enabling their matching across different images. Once a keypoint is detected, a descriptor is computed to create a unique, robust, and compact representation of the image patch around that keypoint. This representation should be distinctive enough to differentiate the keypoint from others, yet invariant or robust to common image transformations such as changes in illumination, viewpoint, scale, and rotation. This section reviews hand-crafted and learned descriptors in chronological order, focusing on their description mechanisms.

SIFT Descriptor (2004):

The SIFT descriptor represents one of the most sophisticated and influential feature description methods in computer vision (Lowe, 2004). The descriptor generates a 128-dimensional floating-point vector that demonstrates distinctiveness and robustness across a wide range of imaging conditions, establishing it as a gold standard for feature description in challenging real-world applications.

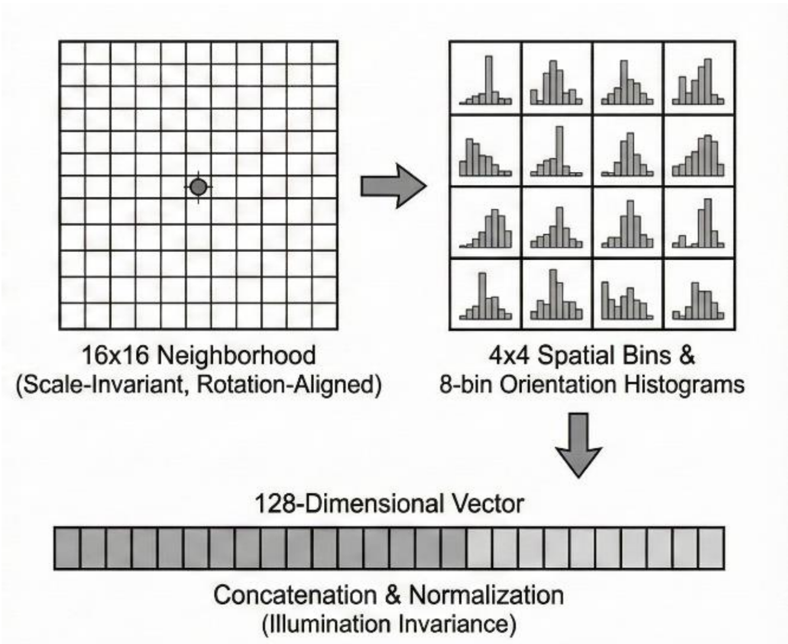


Figure 2.13 SIFT Descriptor Concept

The descriptor construction process begins after keypoint detection and orientation assignment, ensuring that the resulting feature description is both rotation-invariant and anchored to a stable local coordinate system. SIFT examines a 16×16 pixel neighborhood around each keypoint, carefully scaled according to the keypoint's characteristic scale to ensure scale invariance. This region is systematically divided into a 4×4 grid of sub-regions, creating 16 spatial bins that capture the local spatial distribution of gradient information.

Within each of the 16 sub-regions, SIFT computes a gradient orientation histogram containing 8 directional bins spanning the full 360-degree range. This histogram captures the predominant edge orientations within that spatial region, providing information about local structural patterns. The gradient magnitudes are weighted by a Gaussian function centered at the keypoint location, ensuring that pixels closer to the keypoint center contribute more significantly to the descriptor.

The algorithm incorporates normalization procedures to achieve illumination invariance. After concatenating the 16 orientation histograms into a single 128-dimensional vector, SIFT applies unit vector normalization to eliminate the effects of overall brightness variations. Subsequently, the method clips values exceeding 0.2 and renormalizes the vector, reducing the impact of large gradient magnitudes that might arise from illumination changes.

In UAV applications, SIFT descriptors provide stability for scenarios involving significant environmental variations, altitude changes, and lighting conditions that commonly occur in aerial imaging. The descriptor's ability to maintain distinctiveness across different seasons, weather conditions, and times of day makes it particularly valuable for long-term mapping and localization applications.

BRIEF Descriptor (2010):

Binary Robust Independent Elementary Features (BRIEF) introduced a paradigm shift in feature description by demonstrating that effective descriptors could be constructed using simple binary tests rather than complex gradient calculations (Calonder et al., 2010). BRIEF emerged from the recognition that many computer vision applications required fast descriptor computation and matching, even if this meant accepting some reduction in discrimination ability compared to more sophisticated floating-point descriptors.

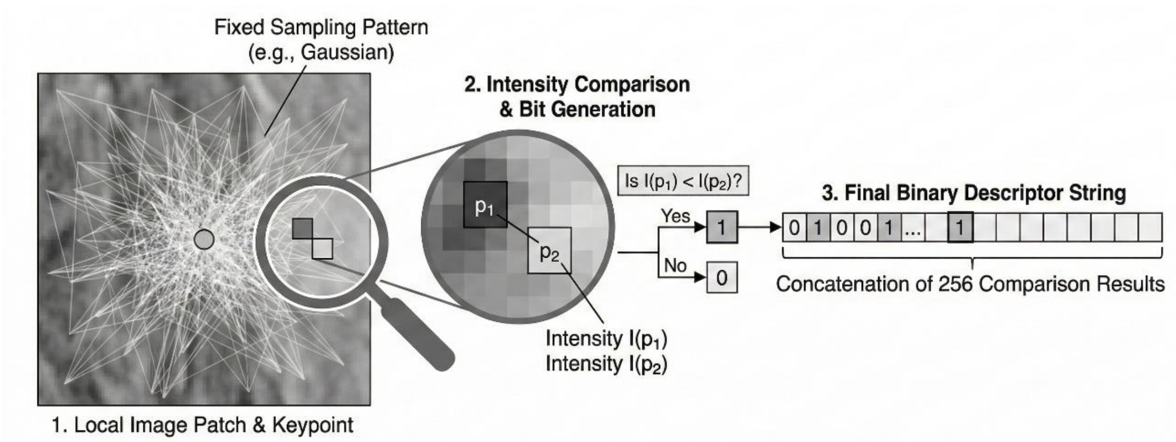


Figure 2.14 BRIEF Descriptor Construction

The algorithm's innovation lies in its simplicity: BRIEF generates compact binary descriptors by performing straightforward intensity comparisons between carefully selected pairs of pixels within the local neighborhood of a keypoint. The method typically employs 256 pixel pairs, chosen according to a predetermined spatial pattern within a square patch around the keypoint. For each pixel pair, BRIEF assigns a binary value based on which pixel exhibits higher intensity, creating a single bit of information that captures a specific aspect of the local intensity distribution.

The selection of pixel pair locations follows statistically optimized patterns designed to maximize descriptor distinctiveness while maintaining computational efficiency. Research has explored various spatial distributions including uniform random sampling, Gaussian-distributed sampling, and structured grid patterns. The choice of sampling pattern significantly impacts descriptor quality, with Gaussian distributions around the keypoint center often providing improved results by concentrating tests in the most stable regions.

BRIEF's binary representation offers computational advantages for both descriptor computation and matching operations. The descriptor calculation requires only simple intensity lookups and comparisons, avoiding the floating-point gradient computations and histogram accumulations required by methods like SIFT. Matching between BRIEF descriptors employs the Hamming distance, which can be computed using bit-wise XOR operations followed by bit counting, enabling rapid similarity assessment.

The descriptor's compact 256-bit representation provides memory advantages compared to floating-point alternatives. This compactness proves particularly valuable in applications involving large-scale feature databases or memory-constrained environments. Additionally, the binary nature supports specialized data structures and indexing schemes that can accelerate nearest-neighbor search operations.

However, BRIEF's emphasis on computational efficiency introduces certain limitations. The basic algorithm lacks built-in rotation invariance, as the fixed pixel pair patterns become misaligned when images are rotated. Similarly, BRIEF demonstrates sensitivity to scale changes since the fixed patch size may not capture appropriate local structure at different scales. These limitations require additional preprocessing steps or algorithmic modifications to achieve invariance to common geometric transformations.

Despite these constraints, BRIEF's speed and reasonable discrimination ability have made it a popular choice for applications prioritizing computational efficiency. The algorithm has inspired numerous subsequent binary descriptor methods that build upon its fundamental approach while addressing its invariance limitations. In UAV applications, BRIEF's rapid computation makes it suitable for real-time onboard processing scenarios where power consumption and computational resources are constrained, particularly when the imaging conditions provide relatively stable geometric relationships.

DAISY Descriptor (2010):

The DAISY descriptor derives its name from the distinctive flower-like arrangement of local image regions used to construct the histogram-based feature representation (Tola et al., 2010). This descriptor addresses feature description specifically designed for dense correspondence applications requiring descriptors at every pixel location. DAISY emerged from the recognition that many computer vision applications, particularly dense stereo matching, optical flow estimation, and wide-baseline correspondence, require feature descriptions not just at sparse keypoint locations but across dense spatial grids covering entire image regions.

The architecture of DAISY employs a hierarchical spatial organization inspired by biological vision systems and designed for computation across dense spatial arrays. The descriptor computes gradient orientation histograms within circular neighborhoods of increasing radii, creating a multi-scale representation that captures both fine-scale local details and broader spatial context. This concentric circular arrangement provides natural rotation symmetry while enabling efficient computation through convolution operations.

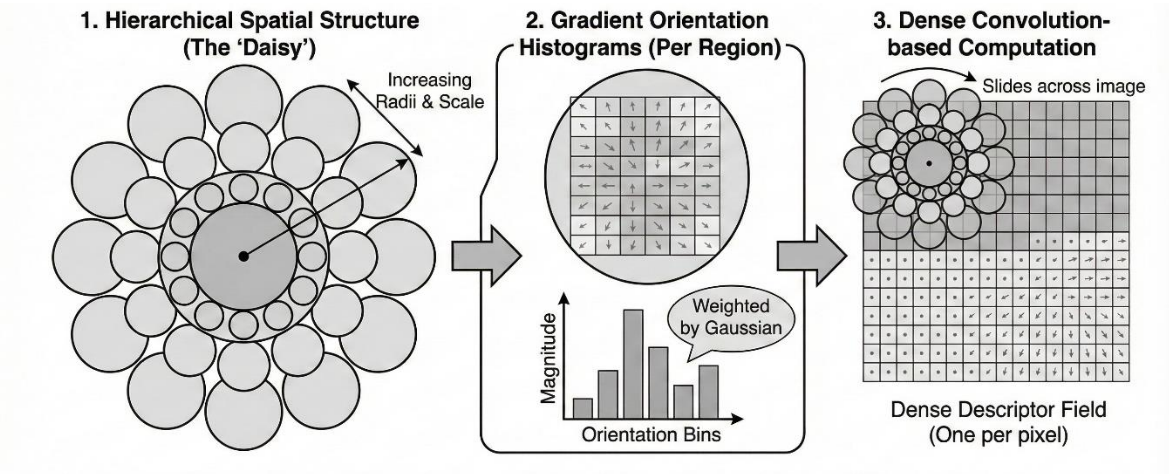


Figure 2.15 Daisy Descriptor Concept

Gradient orientation computation forms the foundation of DAISY's discriminative capability, employing gradient estimation techniques that provide reliable orientation information even in challenging imaging conditions. The algorithm computes gradient magnitude and orientation at each pixel location using derivative operators, constructing orientation histograms that capture the local texture and structural characteristics within different spatial neighborhoods.

The multi-scale spatial organization employs concentric circles of increasing radii around each spatial location, with each circle containing multiple sampling points that contribute to orientation histograms at different scales. This hierarchical arrangement enables the descriptor to capture information at multiple spatial frequencies simultaneously, providing both local detail preservation and broader contextual information that enhances discriminative power.

Convolution-based computation represents a key innovation in DAISY that enables efficient descriptor extraction across dense spatial grids. Rather than computing descriptors independently at each location, DAISY employs convolution operations that efficiently propagate and combine orientation information across spatial neighborhoods. This computational approach reduces the computational complexity compared to naive dense descriptor computation while maintaining accuracy.

Gaussian weighting schemes are employed throughout the DAISY computation to provide smooth spatial falloff characteristics that improve robustness to small spatial perturbations and noise. The Gaussian weighting functions are applied both within individual orientation histogram computations and across the multi-scale spatial organization, creating descriptors with smooth variation characteristics that enhance matching stability.

Rotation handling in DAISY can be achieved through post-processing operations that align descriptors based on dominant orientation estimates or through rotation-invariant histogram organization schemes. While the basic DAISY formulation is not naturally rotation invariant, various extensions and modifications have been developed to provide rotation invariance when required for specific applications.

Dense matching represents the primary application domain for DAISY descriptors, where the algorithm's efficient computation and discriminative characteristics enable correspondence establishment across large spatial regions. The descriptor's design specifically addresses the computational challenges and accuracy requirements of dense matching applications.

Performance characteristics demonstrate that DAISY achieves good discriminative power and robustness to photometric variations while maintaining computational efficiency suitable for dense processing requirements. The multi-scale architecture provides reasonable invariance to moderate scale changes and geometric transformations commonly encountered in stereo matching and optical flow applications.

Computational efficiency benefits arise from the convolution-based implementation that supports parallel processing and efficient memory utilization patterns. Modern implementations can leverage specialized hardware acceleration and optimized numerical libraries to achieve real-time performance even for high-resolution dense processing applications.

For UAV applications, DAISY descriptors prove particularly valuable in scenarios requiring dense correspondence estimation, such as detailed terrain reconstruction, precise navigation, obstacle avoidance systems, and high-resolution mapping applications. The descriptor's ability to provide correspondences across dense spatial grids supports scene understanding and navigation capabilities that require detailed spatial information beyond sparse keypoint matching.

ORB Descriptor (2011):

The ORB descriptor represents an enhancement of the BRIEF descriptor that addresses its rotation invariance limitation while maintaining the computational efficiency advantages of binary feature description (Rublee et al., 2011). ORB emerged as a complete patent-free solution that combines effective feature detection with robust description, providing the computer vision community with a high-performance alternative to proprietary algorithms.

The descriptor construction process builds directly upon BRIEF's proven binary comparison framework while incorporating orientation compensation to achieve rotation invariance. After the ORB detector identifies keypoints and computes their dominant orientations using the intensity centroid method, the descriptor algorithm rotates the standard BRIEF sampling pattern according to each keypoint's canonical orientation. This rotation aligns the descriptor's spatial tests with the keypoint's local coordinate system, ensuring that identical features detected at different rotations produce similar descriptor values.

The pattern rotation process involves applying a rotation matrix to the predetermined pixel pair coordinates, transforming the fixed BRIEF pattern into an orientation-normalized configuration. This geometric transformation ensures that the binary tests capture consistent spatial relationships regardless of the image's overall orientation. The rotation computation utilizes precomputed lookup tables for common rotation angles to maintain computational efficiency, particularly important for real-time applications.

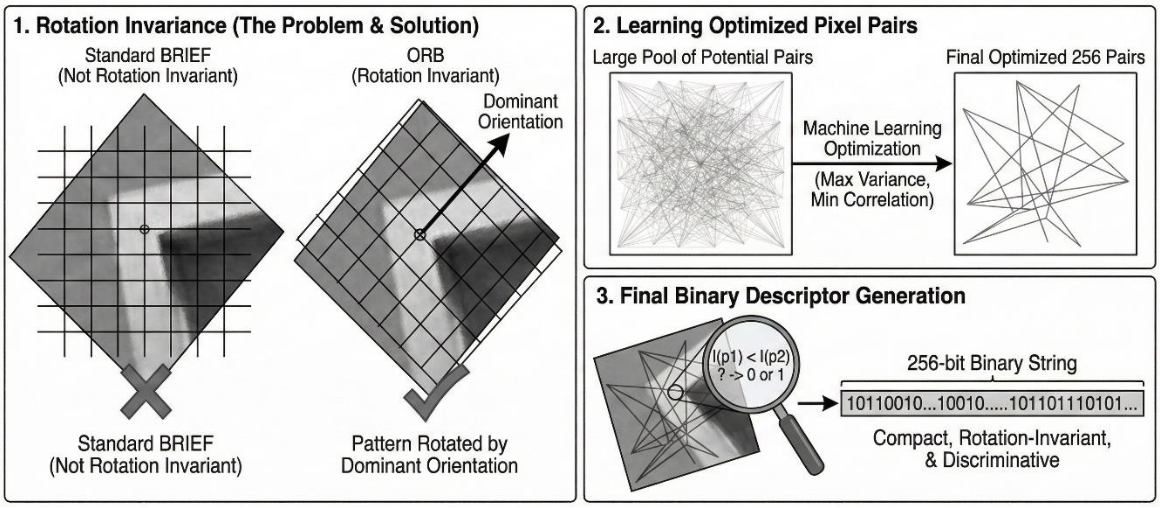


Figure 2.16 ORB Descriptor Concept

ORB incorporates a learning mechanism to optimize the selection of pixel pairs for descriptor construction. Rather than using randomly distributed or heuristically designed patterns, ORB employs a systematic training process that identifies the most discriminative pixel pair combinations. This optimization analyzes large datasets to determine which spatial relationships provide maximum distinctiveness while maintaining low correlation between different bits of the descriptor.

The learning process focuses on two objectives: maximizing the variance of each bit position across different image patches to ensure high discriminative power, and minimizing correlations between different bit positions to reduce redundancy in the descriptor representation. This optimization results in a more efficient use of the available 256 bits, improving matching performance compared to non-optimized binary descriptors.

ORB maintains the computational efficiency advantages of binary descriptors while significantly improving results compared to basic BRIEF. The descriptor calculation requires only slightly more computation than BRIEF due to the orientation compensation step, but still operates much faster than floating-point alternatives. Matching continues to employ Hamming distance calculations, enabling rapid similarity assessment necessary for real-time applications.

The algorithm demonstrates good performance across a wide range of practical scenarios, particularly excelling in applications involving moderate geometric transformations and varying imaging conditions. ORB's combination of rotation invariance, computational efficiency, and patent-free availability has made it one of the most widely adopted feature description methods in practical computer vision systems.

In UAV applications, ORB descriptors provide an appropriate balance between computational efficiency and robustness, making them particularly suitable for onboard processing scenarios where both performance and power consumption are important considerations. The descriptor's rotation invariance proves especially valuable for aerial platforms where camera orientation frequently varies due to vehicle dynamics or intentional gimbal movements.

BRISK Descriptor (2011):

The BRISK descriptor represents a binary descriptor that achieves computational efficiency while maintaining reasonable matching quality across diverse imaging conditions (Leutenegger et al., 2011). The descriptor design philosophy emphasizes the creation of compact binary representations that can be computed rapidly and matched using simple bit-wise operations, making it particularly suitable for real-time applications and resource-constrained environments.

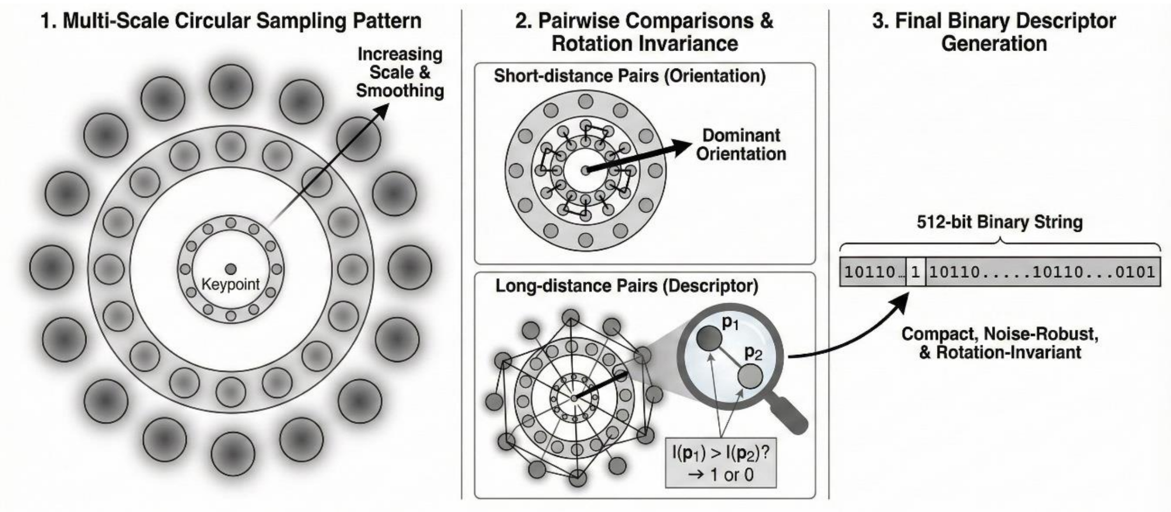


Figure 2.17 BRISK Descriptor Concept

The descriptor construction process employs a carefully designed circular sampling pattern that captures local intensity relationships around detected keypoints. This pattern consists of multiple concentric circles with positioned sampling points, creating a hierarchical representation that captures both fine-scale and coarse-scale spatial relationships. The multi-ring structure enables the descriptor to encode information about local intensity distributions while maintaining computational efficiency.

The sampling pattern incorporates points at varying distances from the keypoint center, enabling the capture of both local detail and broader contextual information. The concentric ring arrangement ensures that the descriptor can adapt to different feature scales and provides robustness against minor localization errors. Each sampling point undergoes Gaussian smoothing with a kernel size appropriate to its distance from the keypoint center, reducing noise sensitivity while preserving essential structural information.

BRISK's binary descriptor generation process involves systematic pairwise comparisons between carefully selected sampling points. The algorithm employs two distinct types of comparisons: short-distance pairs that capture fine-scale intensity patterns and provide rotation estimation capabilities, and long-distance pairs that encode broader spatial relationships for descriptor construction. This dual-purpose design enables simultaneous orientation assignment and descriptor computation within a unified framework.

The rotation invariance mechanism analyzes the short-distance pairs to estimate the keypoint's canonical orientation. The algorithm computes local gradients between nearby sampling points and constructs orientation histograms that reveal predominant directional patterns. The dominant orientation is used to rotate the long-distance pair pattern, ensuring that descriptor values remain consistent regardless of image rotation.

The binary descriptor generation employs the rotated long-distance pairs to create a 512-bit binary string through intensity comparisons. Each bit represents the result of comparing intensity values between a specific pair of sampling points, with the bit set to one if the first point is brighter than the second, and zero otherwise. This simple comparison rule generates a compact binary representation that encodes essential local structure information.

BRISK incorporates noise reduction strategies throughout the descriptor computation process. The Gaussian smoothing applied to sampling points reduces the impact of pixel-level noise, while the statistical nature of multiple pairwise comparisons provides natural robustness against local intensity variations. The binary representation itself contributes to noise robustness by quantizing continuous intensity relationships into discrete binary decisions.

The resulting 512-bit descriptor provides a balance between compactness and discriminative power. The binary representation enables matching using Hamming distance calculations, which can be computed using fast bit-wise XOR operations followed by population counting. This computational efficiency makes BRISK descriptors particularly suitable for applications requiring rapid similarity assessment across large feature databases.

In UAV applications, BRISK descriptors provide robust performance for scenarios involving moderate scale and rotation variations while maintaining the computational efficiency essential for real-time onboard processing. The descriptor's noise robustness proves valuable for aerial imagery that may be affected by platform vibrations, atmospheric disturbances, or varying image quality conditions commonly encountered in UAV operations.

FREAK Descriptor (2012):

Fast Retina Keypoint (FREAK) represents a biologically-inspired approach to feature description that draws design principles from the human visual system, specifically mimicking the retinal sampling patterns found in biological vision systems (Alahi et al., 2012). This descriptor emerged from neuroscience research demonstrating that the human retina employs non-uniform spatial sampling with higher resolution at the center and progressively lower resolution toward the periphery, a pattern that proves efficient for visual information processing.

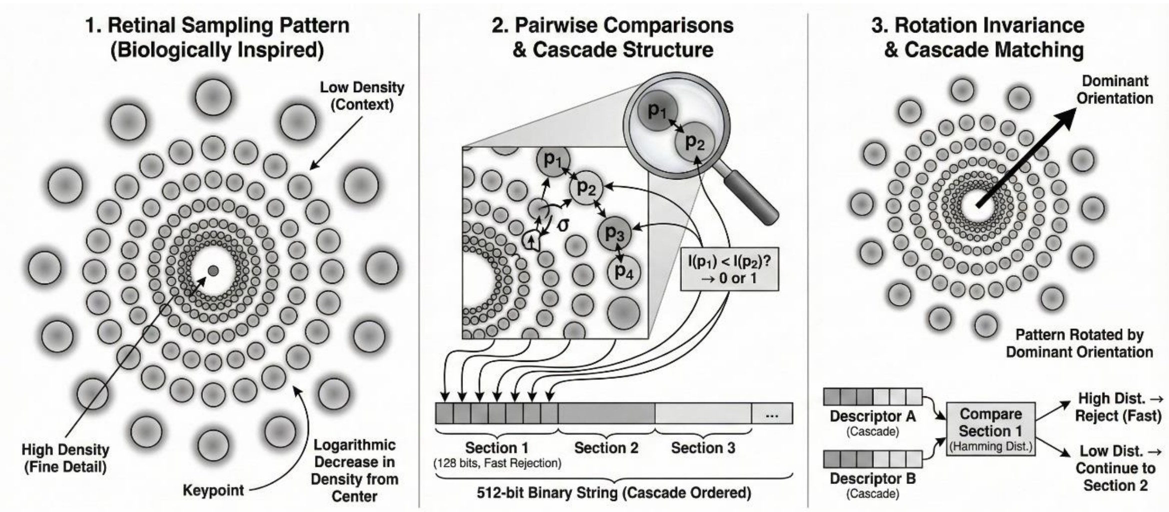


Figure 2.18 FREAK Descriptor Concept

The algorithm's core innovation lies in its retinal-inspired sampling pattern that mimics the distribution of photoreceptors in biological vision systems. The sampling pattern employs concentric circles around the keypoint center, with point density decreasing logarithmically from center to periphery. This arrangement concentrates computational resources on the most informative central region while still capturing important contextual information from the surrounding area, creating a balance between detail preservation and computational efficiency.

The multi-scale sampling structure consists of multiple concentric rings with decreasing point densities, creating a hierarchical representation that captures information at different spatial frequencies. The central region contains densely packed sampling points that capture fine-scale details and texture information, while outer rings contain more sparsely distributed points that encode broader spatial relationships and contextual patterns. This hierarchical approach enables the descriptor to adapt to features of varying scales and complexities.

FREAK employs a cascade matching approach inspired by biological visual processing mechanisms. The descriptor is structured to enable coarse-to-fine matching, where initial comparisons use a subset of the most discriminative bits to rapidly eliminate obviously non-matching candidates. Only promising candidates undergo complete descriptor comparison, reducing computational overhead in matching operations while maintaining reasonable accuracy.

The binary descriptor generation process involves systematic pairwise intensity comparisons between selected sampling points. The selection of comparison pairs follows optimization procedures that maximize discriminative power while ensuring decorrelation between different descriptor bits. This optimization process analyzes large training datasets to identify the most informative spatial relationships, resulting in efficient use of the available descriptor bits.

Rotation invariance is achieved through orientation estimation based on local gradient analysis within the sampling pattern. The algorithm computes gradient directions between nearby sampling points and constructs orientation histograms that reveal the dominant local orientation. The estimated canonical orientation is used to rotate the comparison pattern, ensuring consistent descriptor values regardless of image rotation.

The cascade structure enables early rejection of non-matching descriptors, improving computational efficiency in large-scale matching scenarios. The algorithm arranges descriptor bits in order of discriminative power, allowing matching algorithms to make confident rejection decisions using only the first several bits for obviously dissimilar features. This early termination capability proves particularly valuable in applications involving large feature databases or real-time matching requirements.

FREAK's 512-bit binary representation provides good discriminative power while maintaining the computational efficiency advantages of binary descriptors. The biologically-inspired design principles result in descriptors that demonstrate good robustness to noise, illumination changes, and geometric transformations commonly encountered in real-world imaging scenarios.

In UAV applications, FREAK's hierarchical sampling pattern and cascade matching capabilities make it particularly suitable for scenarios involving large image databases or applications requiring efficient feature matching across extensive aerial surveys. The descriptor's robustness to imaging variations proves valuable for long-term monitoring applications where consistent feature matching across different seasons, lighting conditions, and atmospheric effects is essential.

KAZE Descriptor (2012):

The KAZE descriptor represents a paradigm shift in feature description methodology by operating exclusively within the nonlinear scale-space framework established by the KAZE feature detection algorithm (Alcantarilla et al., 2012). This descriptor differs from traditional approaches by maintaining consistency with the nonlinear diffusion process that defines the scale-space representation, ensuring that both detection and description operate under the same mathematical framework and preserving the enhanced structural coherence achieved through nonlinear diffusion filtering.

The descriptor construction process operates directly on the gradient information derived from the nonlinear scale-space representation computed during the feature detection phase. Unlike conventional descriptors that work on linearly filtered images, the KAZE descriptor leverages the edge-preserving properties of nonlinear diffusion to compute more accurate and robust gradient estimates. The nonlinear diffusion process selectively smooths homogeneous regions while preserving important edge and corner information, resulting in gradient fields that better reflect the true underlying image structure.

The spatial organization follows established principles from SIFT-like descriptors but adapts these concepts to work effectively within the nonlinear scale-space framework. The algorithm divides the keypoint neighborhood into subregions and computes gradient orientation histograms within each subregion, capturing the local intensity distribution patterns that characterize the feature. However, the gradient computations benefit from the enhanced structural preservation achieved through nonlinear diffusion, potentially providing more accurate local orientation estimates.

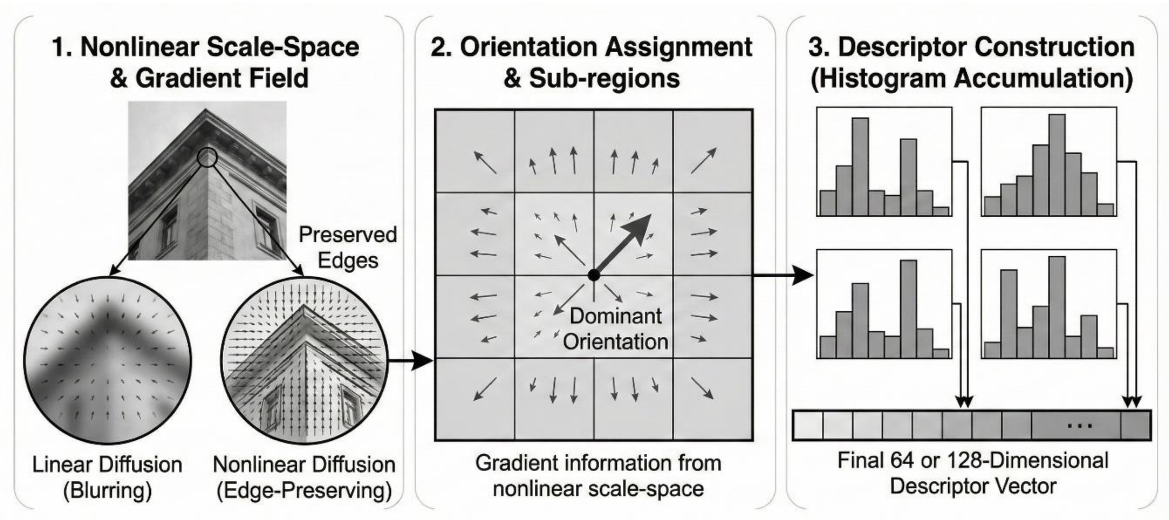


Figure 2.19 KAZE Descriptor Concept

Scale normalization in the KAZE descriptor framework presents unique challenges and opportunities compared to traditional approaches. The nonlinear scale-space representation provides scale information that naturally preserves important structural details across different scales, potentially offering more robust scale normalization. The descriptor sampling window size is determined based on the characteristic scale identified during detection, ensuring appropriate spatial coverage for features detected at different scale levels.

Orientation assignment follows modified SIFT-like procedures but operates on gradient information derived from the nonlinear scale-space representation. The algorithm computes gradient magnitude and orientation at sampling points within the keypoint neighborhood, constructing orientation histograms that capture the dominant local gradient directions. The enhanced edge preservation properties of nonlinear diffusion potentially provide more accurate orientation estimates, particularly for features associated with significant structural boundaries.

The descriptor vector construction involves systematic accumulation of gradient magnitude and orientation information across the spatial subregions surrounding the keypoint. Each subregion contributes weighted gradient statistics to the final descriptor vector, with weighting schemes designed to emphasize central regions while including relevant contextual information from peripheral areas. The resulting descriptor captures both local intensity patterns and broader spatial relationships within the feature neighborhood.

Robustness characteristics of the KAZE descriptor derive directly from the properties of the underlying nonlinear diffusion process. The edge-preserving nature of nonlinear diffusion helps maintain consistent descriptor values across different imaging conditions, particularly when images exhibit varying noise levels or different acquisition parameters. The structural preservation capabilities potentially provide better invariance to illumination changes and mild geometric transformations.

The computational complexity of KAZE descriptor extraction reflects the mathematical framework underlying the nonlinear diffusion process. While more computationally intensive than basic binary descriptors, the enhanced structural preservation and potential accuracy improvements may justify the additional computational cost in applications where feature matching quality is important.

For UAV applications, the KAZE descriptor's emphasis on structural preservation proves particularly valuable in scenarios involving complex terrain features, urban environments with significant edge content, or applications requiring precise feature localization across different imaging conditions. The enhanced gradient estimation capabilities may provide improved matching accuracy for features associated with important structural elements in aerial imagery.

LUCID Descriptor (2012):

Locally Uniform Comparison Image Descriptor (LUCID) represents a different approach to feature description that prioritizes computational simplicity while maintaining adequate discriminative power for resource-constrained applications (Ziegler et al., 2012). This descriptor emerged from the recognition that many computer vision applications, particularly those deployed on mobile devices, embedded systems, or real-time processing scenarios, require feature matching capabilities that operate within strict computational and energy constraints.

The core innovation of LUCID lies in its deliberate simplification of the feature description process, reducing the computational complexity to basic pixel intensity comparisons that can be efficiently implemented on limited-resource hardware platforms. Unlike descriptors that employ complex mathematical transformations or multi-scale analysis, LUCID operates directly on raw pixel intensities using simple comparison operations that require minimal computational overhead and can be vectorized on modern processors.

The algorithm's spatial sampling strategy employs small, fixed-size kernel patterns centered on detected keypoints, with the exact pattern design optimized for capturing essential local information while minimizing computational requirements. The kernel typically consists of a modest number of sampling points arranged in geometric patterns that provide good spatial coverage of the immediate keypoint neighborhood. This limited spatial extent ensures that descriptor computation remains practical while capturing sufficient local information for discrimination.

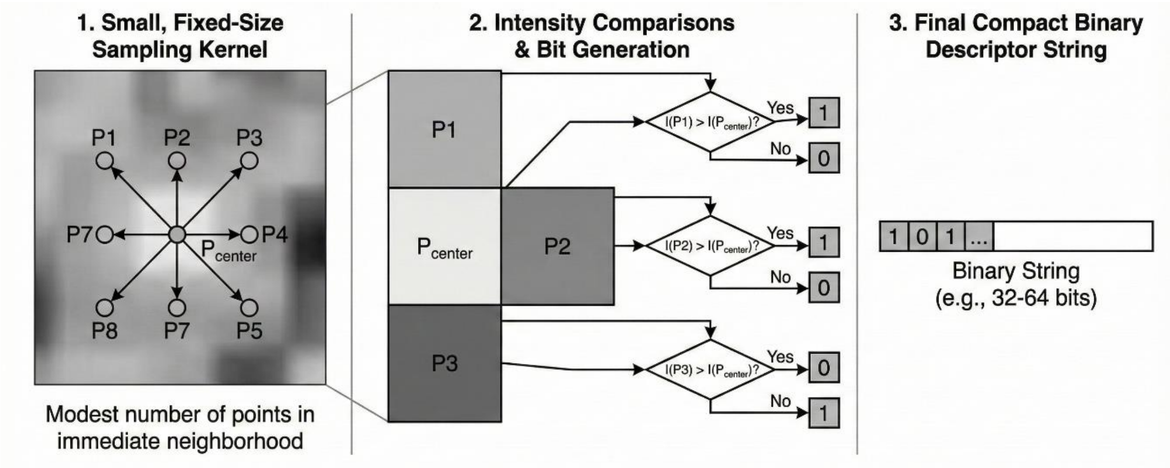


Figure 2.20 LUCID Descriptor Concept

The binary descriptor generation process involves systematic intensity comparisons between selected pixel pairs within the sampling kernel. Each comparison generates a single bit of the descriptor vector, with the comparison result determined by simple intensity threshold operations. The central pixel often serves as a reference point for comparisons, with surrounding pixels compared against this central intensity value to generate binary comparison results.

Pattern selection and optimization represent aspects of LUCID's design philosophy, as the choice of comparison patterns directly impacts both discriminative power and computational efficiency. The algorithm employs patterns that have been empirically validated to provide good feature discrimination while maintaining the simplicity necessary for efficient implementation. These patterns typically focus on capturing local intensity gradients and texture variations that characterize distinctive image regions.

Rotation and scale invariance in LUCID present challenges given the descriptor's emphasis on simplicity. While the basic algorithm may not incorporate sophisticated geometric normalization procedures, variations of LUCID can employ simple orientation estimation techniques or multi-scale sampling strategies to improve robustness to geometric transformations. However, these enhancements must be carefully balanced against the design goal of computational simplicity.

The descriptor's compact binary representation enables efficient storage and fast matching operations using specialized bit manipulation techniques. Hamming distance calculations between LUCID descriptors can be performed using bit-wise operations, making the matching process fast compared to more complex descriptor types. This efficiency proves particularly valuable in applications requiring real-time performance or processing large numbers of features.

Performance characteristics of LUCID reflect the trade-offs involved in prioritizing efficiency over sophistication. While the descriptor may not achieve the discriminative power of more complex alternatives, it provides adequate performance for many practical applications where computational constraints are important. The simplicity also contributes to robustness in scenarios where more complex algorithms might fail due to implementation constraints or numerical precision limitations.

In UAV applications, LUCID's efficiency makes it particularly suitable for onboard processing scenarios where computational resources are limited, power consumption is important, or real-time processing requirements must be met. The descriptor's simplicity can be advantageous for applications involving continuous monitoring, where processing efficiency directly impacts mission duration and operational effectiveness.

AKAZE Descriptor (2013):

Accelerated-KAZE (AKAZE) descriptor represents an evolutionary advancement in nonlinear scale-space feature description that addresses the computational limitations of the original KAZE descriptor while preserving its fundamental advantages in structural preservation and robustness (Alcantarilla et al., 2013). This descriptor emerges from careful analysis of the computational bottlenecks in nonlinear diffusion processing and implements strategic algorithmic modifications that dramatically improve processing speed without sacrificing the essential benefits of nonlinear scale-space representation.

The fundamental innovation underlying the AKAZE descriptor lies in its adoption of Fast Explicit Diffusion (FED) schemes as a replacement for the computationally intensive Additive Operator Splitting (AOS) methods employed in the original KAZE algorithm. FED schemes provide mathematically equivalent solutions to the nonlinear diffusion equations while enabling much more efficient numerical implementation through explicit time-stepping procedures that can be optimized for modern processor architectures.

The Modified Local Difference Binary (M-LDB) pattern serves as the core descriptor generation mechanism, representing a sophisticated adaptation of Local Binary Pattern (LBP) concepts to work effectively within the nonlinear scale-space framework. M-LDB patterns capture local intensity and gradient relationships through systematic comparisons of carefully selected spatial locations within the keypoint neighborhood, generating binary descriptor bits that encode essential structural information preserved by the nonlinear diffusion process.

Scale-space construction in AKAZE employs the same fundamental mathematical principles as KAZE but implements these through more efficient computational procedures. The algorithm constructs nonlinear scale pyramids using FED-based diffusion that preserves edge information while smoothing noise and irrelevant details. This scale-space representation provides the foundation for robust feature description that maintains structural coherence across different scales and imaging conditions.

The binary descriptor generation process systematically samples gradient and intensity information within the nonlinear scale-space representation, comparing values at strategically selected spatial locations to generate discriminative binary patterns. The comparison operations focus on relationships that capture essential structural characteristics while being robust to noise and illumination variations. The resulting binary vectors encode local image structure in a compact format suitable for efficient matching operations.

Orientation estimation and normalization follow established principles adapted for the nonlinear scale-space context. The algorithm computes dominant orientations based on gradient information derived from the edge-preserving diffusion process, potentially providing more accurate orientation estimates than conventional linear filtering approaches. The orientation normalization ensures rotational invariance while preserving the structural advantages of nonlinear processing.

The computational efficiency improvements achieved through FED-based implementation enable AKAZE descriptors to approach the processing speeds of conventional linear-filtering methods while maintaining the superior structural preservation capabilities of nonlinear approaches. This efficiency breakthrough makes nonlinear scale-space methods practical for applications requiring real-time processing or handling large numbers of features.

Robustness characteristics inherit the fundamental advantages of nonlinear diffusion processing, including enhanced noise tolerance, better preservation of structural boundaries, and improved stability under varying imaging conditions. The edge-preserving properties of the underlying diffusion process contribute to descriptor consistency across different image acquisition parameters and environmental conditions.

The binary nature of AKAZE descriptors enables efficient matching using Hamming distance calculations, combining the speed advantages of binary operations with the structural preservation benefits of nonlinear scale-space processing. This combination proves particularly valuable in applications requiring both high matching accuracy and computational efficiency.

For UAV applications, AKAZE descriptors offer an excellent balance between computational efficiency and matching performance, making them suitable for onboard processing scenarios where both accuracy and speed are important. The enhanced structural preservation capabilities prove particularly valuable for terrain analysis, infrastructure monitoring, and precision navigation applications where consistent feature matching across varying conditions is essential.

BOOST Descriptor (2013):

BoostDesc, commonly referred to as BOOST descriptor, represents a machine learning-driven approach to feature description that fundamentally departs from traditional hand-crafted descriptor design methodologies (Trzcinski et al., 2013). This innovative descriptor leverages advanced boosting algorithms to systematically learn optimal feature description strategies from large training datasets, automatically discovering discriminative patterns and relationships that might not be apparent through manual algorithm design.

The fundamental paradigm underlying BOOST descriptors involves the application of boosting machine learning techniques to automatically select and combine simple binary tests into highly discriminative feature descriptors. Rather than relying on predetermined mathematical formulations or geometric assumptions, the algorithm learns to identify the most informative local image characteristics through systematic analysis of training data containing diverse feature types and matching scenarios.

The weak learner selection process forms the core of the BOOST descriptor methodology, where individual binary tests serve as fundamental building blocks for descriptor construction. These weak learners typically consist of simple intensity comparison operations, local gradient evaluations, or basic pattern matching tests that can be efficiently computed but individually provide limited discriminative power. The boosting algorithm systematically evaluates vast numbers of potential weak learners to identify those that contribute most effectively to feature discrimination.

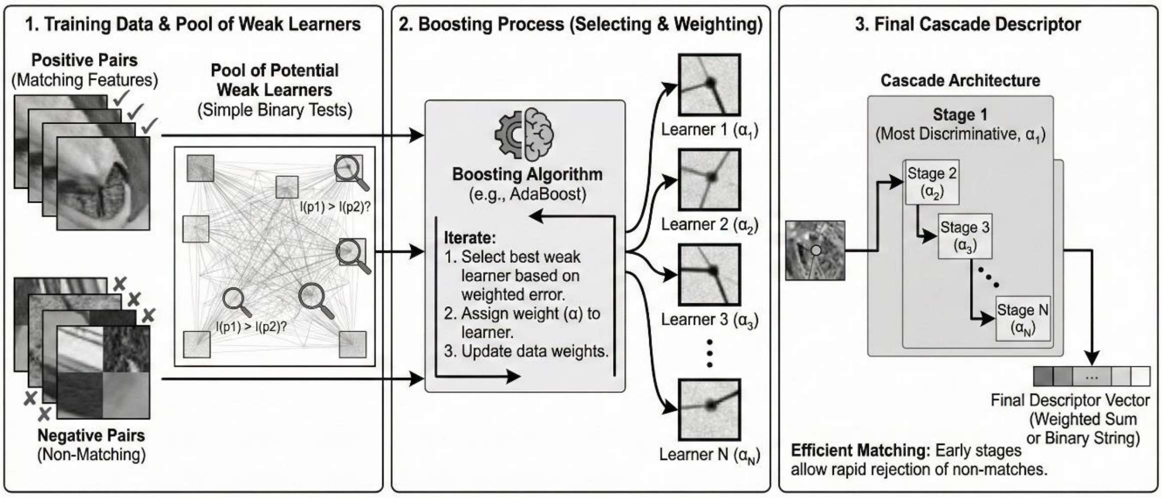


Figure 2.21 BOOST Descriptor Concept

Cascade architecture implementation enables efficient descriptor construction and matching through hierarchical organization of the selected weak learners. The boosting process arranges learned tests in order of discriminative importance, creating a cascade structure where early stages can rapidly eliminate obviously non-matching feature pairs while later stages provide refined discrimination for challenging cases. This cascade approach significantly improves computational efficiency in large-scale matching applications.

Training data requirements for BOOST descriptors necessitate carefully curated datasets containing diverse feature types, imaging conditions, and geometric transformations representative of target applications. The machine learning process analyzes positive and negative matching examples to learn discriminative patterns, requiring substantial computational resources during the training phase but resulting in highly optimized descriptors for specific application domains.

The flexibility of the BOOST framework enables generation of both binary and floating-point descriptor variants, allowing optimization for different application requirements. Binary versions prioritize computational efficiency and storage minimization, while floating-point variants can achieve higher discriminative power when computational resources permit. This adaptability makes BOOST descriptors suitable for diverse deployment scenarios ranging from resource-constrained mobile applications to high-performance computational platforms.

Performance characteristics of BOOST descriptors reflect the advantages of data-driven optimization, often achieving superior discriminative power compared to hand-crafted alternatives on datasets similar to the training data. However, performance may degrade when applied to imaging conditions or feature types significantly different from the training distribution, highlighting the importance of comprehensive training data selection.

Computational efficiency during deployment benefits from the optimized cascade structure, enabling early termination for non-matching descriptors while providing detailed analysis for promising candidates. The learned weak learners are typically simple operations that can be efficiently implemented on various hardware platforms, making BOOST descriptors practical for real-time applications despite their sophisticated underlying machine learning foundation.

Generalization capabilities depend critically on the diversity and representativeness of training data, with well-trained BOOST descriptors demonstrating good robustness to imaging variations within the training distribution. The machine learning approach can potentially discover subtle discriminative patterns that human designers might overlook, leading to improved matching performance in specific application domains.

For UAV applications, BOOST descriptors offer the potential for domain-specific optimization through training on aerial imagery datasets, potentially achieving superior performance compared to generic descriptors. The ability to tailor descriptor characteristics to specific UAV mission requirements, imaging conditions, or terrain types provides opportunities for significant performance improvements in specialized applications.

VGG Descriptor (2014):

The VGG descriptor implementation in OpenCV represents a significant paradigm shift in feature description methodology, fundamentally departing from traditional hand-crafted approaches by employing deep learning architectures to automatically learn hierarchical feature representations from vast training datasets. This descriptor draws inspiration from innovative research that demonstrated how Convolutional Neural Networks (CNNs) could automatically discover highly effective local feature representations through systematic training on large-scale image datasets, achieving superior performance compared to manually designed descriptor algorithms (Simonyan et al., 2014).

The underlying architecture leverages pre-trained VGG network models that have been trained on extensive image classification tasks, inheriting sophisticated feature extraction capabilities that have been optimized through gradient-based learning processes. The VGG network architecture employs deep convolutional structures with small receptive fields and multiple layers, creating hierarchical representations that capture increasingly complex visual patterns from low-level edges and textures to high-level semantic concepts.

Feature extraction methodology involves processing normalized image patches centered on detected keypoints through the pre-trained VGG network, extracting activation vectors from intermediate network layers that serve as the descriptor representation. This process differs fundamentally from traditional approaches by utilizing learned filters and nonlinear transformations rather than predetermined mathematical operations, potentially capturing more sophisticated patterns and relationships within local image regions.

The hierarchical nature of deep network representations enables VGG descriptors to encode multiple levels of visual information simultaneously, from fine-scale textural details captured in early network layers to broader structural patterns and semantic relationships encoded in deeper layers. This multi-scale representation capability provides natural robustness to various geometric and photometric transformations that commonly affect traditional hand-crafted descriptors.

Pre-training advantages ensure that VGG descriptors benefit from knowledge acquired through exposure to millions of diverse training images, potentially providing better generalization capabilities compared to descriptors designed through analysis of limited datasets. The extensive training process enables the network to automatically discover optimal filter configurations and feature combinations that might not be apparent through manual design processes.

Computational considerations for VGG descriptors involve trade-offs between descriptor quality and processing efficiency, as deep network inference typically requires more computational resources than traditional descriptor extraction methods. However, modern GPU implementations and optimized inference frameworks can significantly accelerate processing, making VGG descriptors practical for applications where matching accuracy is prioritized over computational simplicity.

Descriptor dimensionality in VGG implementations tends to be higher than traditional approaches, reflecting the rich representational capacity of deep network features. While this increased dimensionality provides enhanced discriminative power, it also impacts storage requirements and matching computational costs, necessitating careful consideration of application-specific constraints.

Robustness characteristics of VGG descriptors often demonstrate superior performance under challenging imaging conditions, illumination variations, and geometric transformations compared to hand-crafted alternatives. The learned features can exhibit better invariance properties that emerge naturally from the training process rather than being explicitly engineered into the algorithm design.

Transfer learning capabilities enable VGG descriptors to leverage representations learned from general image classification tasks for specialized feature matching applications, potentially providing good performance even in domains not specifically represented in the original training data. This transferability represents a significant advantage over domain-specific hand-crafted approaches.

For UAV applications, VGG descriptors offer potential advantages in scenarios involving complex terrain features, varying environmental conditions, or applications requiring high matching accuracy. The learned representations may better capture the diverse visual patterns encountered in aerial imagery, although computational requirements must be balanced against onboard processing constraints and real-time performance requirements.

LATCH Descriptor (2016):

Learned Arrangements of Three Patch Codes (LATCH) represents an innovative advancement in binary descriptor design that fundamentally reconsiders the traditional pairwise comparison paradigm by introducing triplet-based comparison schemes that capture more sophisticated spatial relationships within local image neighborhoods (Levi & Hassner, 2016). This descriptor emerges from the recognition that simple pairwise intensity comparisons, while computationally efficient, may overlook important structural information that could significantly enhance feature discrimination capabilities.

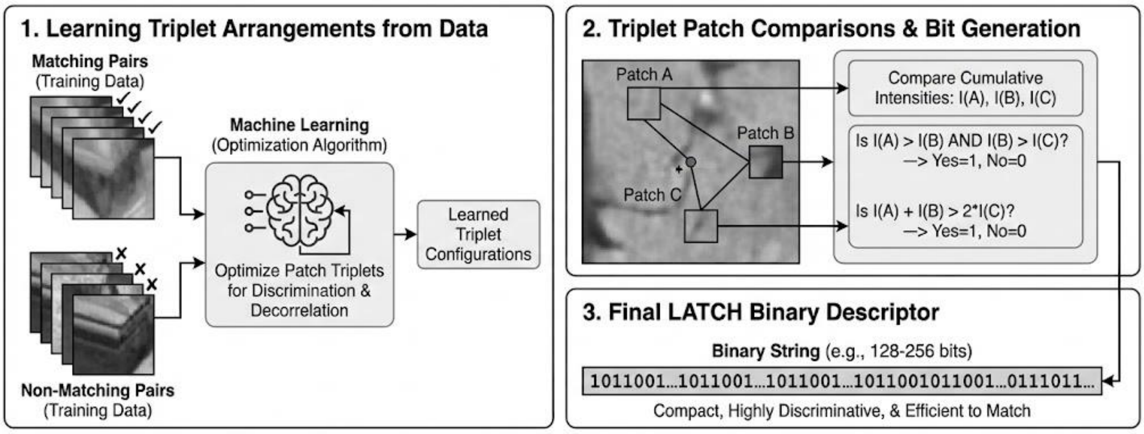


Figure 2.22 LATCH Descriptor Concept

The core innovation underlying LATCH involves the systematic use of triplet patch comparisons rather than conventional pairwise intensity comparisons employed by most binary descriptors. Each triplet consists of three carefully selected spatial locations within the keypoint neighborhood, with the algorithm comparing the cumulative intensity values across these patch regions to generate individual bits of the binary descriptor. This approach captures more complex spatial relationships and intensity distributions that provide enhanced discriminative power.

Patch arrangement strategies in LATCH employ sophisticated learning algorithms to determine optimal spatial configurations for the triplet comparisons. Rather than using predetermined geometric patterns, the algorithm analyzes training datasets to identify patch arrangements that maximize discriminative power while minimizing correlation between different triplet tests. This data-driven optimization process results in patch configurations that are specifically tailored for effective feature discrimination.

The learning methodology underlying patch selection involves extensive analysis of feature matching scenarios to identify triplet arrangements that consistently provide reliable discrimination between matching and non-matching features. The algorithm evaluates numerous potential patch configurations using training datasets containing diverse feature types and imaging conditions, systematically selecting arrangements that demonstrate robust performance across various scenarios.

Spatial encoding efficiency in LATCH benefits from the enhanced information content of triplet comparisons, enabling the generation of highly discriminative binary descriptors with relatively modest bit lengths. Each triplet comparison captures richer spatial information than simple pairwise comparisons, allowing LATCH to achieve superior discrimination performance while maintaining the computational efficiency advantages of binary representations.

The binary descriptor generation process systematically evaluates each learned triplet arrangement within the keypoint neighborhood, computing cumulative intensity values for each patch region and performing comparisons that generate individual descriptor bits. The resulting binary string encodes sophisticated spatial relationships that reflect the optimized patch arrangements determined during the learning phase.

Computational efficiency considerations demonstrate that LATCH maintains the fundamental speed advantages of binary descriptors while providing enhanced discriminative capabilities. The triplet comparisons require only modest additional computational overhead compared to pairwise approaches, making LATCH practical for applications requiring both accuracy and efficiency. The binary nature enables fast matching using Hamming distance calculations.

Robustness characteristics of LATCH benefit from the enhanced spatial information captured through triplet comparisons, providing improved stability under various imaging conditions and geometric transformations. The learned patch arrangements can capture spatial relationships that remain consistent across different viewing conditions, contributing to more reliable feature matching performance.

Performance optimization through machine learning enables LATCH to achieve descriptor characteristics that are specifically optimized for feature matching applications, rather than relying on geometric assumptions or mathematical intuitions that may not reflect optimal discrimination strategies. This data-driven approach can discover effective patch arrangements that might not be apparent through manual design processes.

For UAV applications, LATCH offers an excellent balance between computational efficiency and matching accuracy, making it particularly suitable for scenarios requiring reliable feature matching with moderate computational resources. The enhanced discriminative power provided by triplet comparisons can improve matching reliability in challenging aerial imagery scenarios while maintaining processing speeds suitable for real-time applications.

BEBLID Descriptor (2020):

Boosted Efficient Binary Local Image Descriptor (BEBLID) represents a state-of-the-art advancement in binary descriptor design that leverages sophisticated machine learning optimization techniques to systematically improve the fundamental building blocks of binary feature description (Suárez et al., 2020). This descriptor emerges from comprehensive analysis of the limitations in traditional binary descriptor design and applies advanced boosting algorithms to overcome these constraints through data-driven optimization of pixel comparison patterns.

The fundamental innovation underlying BEBLID involves the application of boosting machine learning techniques to systematically optimize the selection and arrangement of binary tests used in descriptor construction. Unlike traditional approaches that rely on geometric intuition or predetermined patterns, BEBLID employs algorithmic analysis of extensive training datasets to identify pixel comparison configurations that maximize discriminative power while minimizing redundancy between different descriptor components.

Optimization methodology in BEBLID follows sophisticated machine learning protocols that analyze vast numbers of potential binary test configurations to identify those that contribute most effectively to feature discrimination. The boosting algorithm systematically evaluates different pixel pair comparisons, spatial arrangements, and weighting schemes to construct descriptor patterns that are specifically optimized for reliable feature matching across diverse imaging conditions and geometric transformations.

Binary test selection process employs advanced statistical analysis to identify pixel comparison patterns that provide maximum information content while maintaining low correlation between different descriptor bits. This optimization ensures that each bit of the binary descriptor contributes unique discriminative information, maximizing the effectiveness of the limited bit budget available in binary representations.

Training data requirements for BEBLID necessitate carefully curated datasets containing diverse feature types, imaging conditions, and geometric transformations representative of target applications. The machine learning optimization process analyzes positive and negative matching examples to learn optimal comparison patterns, requiring substantial computational resources during the training phase but resulting in highly efficient descriptors for deployment.

The efficiency optimization involved in the boosting approach enables BEBLID to achieve superior matching performance compared to traditional binary descriptors while maintaining the computational efficiency advantages that make binary descriptors attractive for resource-constrained applications. The optimized test selection ensures that computational resources are focused on the most informative comparisons.

Performance characteristics of BEBLID demonstrate significant improvements in matching accuracy compared to earlier binary descriptor designs, reflecting the benefits of data-driven optimization over hand-crafted approaches. The systematic optimization process can discover discriminative patterns that might not be apparent through manual design, resulting in descriptors that better capture the essential characteristics needed for reliable feature matching.

Computational efficiency during deployment remains a primary strength of BEBLID, as the optimized binary tests can be efficiently computed using simple pixel comparison operations that are well-suited for various hardware platforms. The binary nature enables fast matching using specialized bit manipulation techniques and Hamming distance calculations.

Robustness evaluation demonstrates that BEBLID descriptors maintain good performance across varying imaging conditions, geometric transformations, and noise levels within the ranges represented in the training data. The machine learning optimization process can identify comparison patterns that remain stable under various challenging conditions commonly encountered in practical applications.

Generalization capabilities depend on the diversity and representativeness of training datasets, with well-trained BEBLID descriptors demonstrating good performance on feature types and imaging conditions similar to those encountered during training. The optimization process can potentially identify subtle discriminative patterns that provide improved matching reliability in specific application domains.

For UAV applications, BEBLID offers the potential for domain-specific optimization through training on aerial imagery datasets, enabling the development of descriptors that are specifically tailored for the unique characteristics and challenges of UAV-based feature matching. The combination of optimized performance and computational efficiency makes BEBLID particularly suitable for onboard processing scenarios where both accuracy and speed are critical requirements.

TEBLID Descriptor (2021):

Tiny Efficient Binary Local Image Descriptor (TEBLID) represents the culmination of research efforts focused on achieving maximum computational efficiency in binary descriptor design while maintaining competitive discriminative performance for resource-critical applications (Suárez et al., 2021). This descriptor addresses the growing demand for feature matching capabilities in extremely constrained environments such as Internet of Things (IoT) devices, embedded systems, mobile platforms, and edge computing scenarios where computational resources, memory capacity, and energy consumption must be minimized.

The design philosophy underlying TEBLID emphasizes the optimization of every computational aspect of the descriptor extraction and matching process, recognizing that even modest efficiency improvements can have significant cumulative impacts in resource-constrained deployment scenarios. The algorithm represents a careful balance between discriminative power and computational simplicity, employing sophisticated optimization techniques to identify the most efficient possible configurations for binary feature description.

Optimization methodology in TEBLID employs advanced machine learning techniques similar to BEBLID but with additional constraints focused specifically on computational efficiency and memory utilization. The optimization process simultaneously considers multiple performance criteria including descriptor extraction speed, memory requirements, matching computational cost, and discriminative effectiveness, seeking configurations that provide optimal trade-offs across these competing objectives.

Binary test pattern optimization in TEBLID involves systematic analysis of computational complexity for different comparison configurations, prioritizing patterns that can be efficiently implemented on resource-constrained hardware platforms. The algorithm identifies pixel comparison arrangements that minimize memory access patterns, reduce computational dependencies, and enable efficient vectorization or parallel processing when available.

Efficiency enhancement strategies employed in TEBLID include careful optimization of spatial sampling patterns, strategic selection of pixel comparison operations, and intelligent organization of binary tests to enable early termination in matching scenarios. These optimizations collectively contribute to significant reductions in computational overhead while maintaining adequate discriminative capabilities for practical feature matching applications.

Compact representation design ensures that TEBLID descriptors utilize minimal memory resources both for storage and during processing operations. The algorithm employs optimized bit packing techniques and efficient data structures that minimize memory footprint and enable cache-friendly access patterns that are particularly important for embedded system deployments.

Performance validation demonstrates that TEBLID achieves competitive matching accuracy compared to more computationally intensive alternatives while operating within significantly reduced resource constraints. The optimization process identifies discriminative patterns that provide good feature matching performance despite the emphasis on computational efficiency.

Real-time processing capabilities represent a primary strength of TEBLID, enabling feature matching applications that require immediate response times or must process continuous data streams with limited computational resources. The extreme efficiency optimizations make TEBLID suitable for applications where traditional descriptors would be impractical due to resource limitations.

Hardware compatibility considerations ensure that TEBLID can be efficiently implemented across diverse hardware platforms, from high-end processors with specialized vector processing capabilities to simple microcontrollers with limited computational resources. The algorithm design accommodates various hardware architectures while maintaining consistent performance characteristics.

Energy efficiency represents a critical consideration for battery-powered devices and mobile applications, with TEBLID optimizations specifically targeting reduced energy consumption through minimized computational operations and efficient memory access patterns. These optimizations directly translate to extended operational periods for autonomous systems and mobile platforms.

For UAV applications, TEBLID provides opportunities for implementing feature matching capabilities on small, lightweight platforms with strict power and computational constraints. The extreme efficiency optimizations enable continuous monitoring applications, extended mission durations, and deployment of feature matching capabilities on platforms where traditional descriptors would be impractical due to resource limitations. The descriptor's efficiency makes it particularly suitable for swarm robotics applications where multiple UAV platforms must operate within shared computational resource constraints.

2.3. Feature Matching

After features are extracted and described across multiple images, the subsequent critical phase involves establishing reliable correspondences between them through sophisticated matching algorithms. This process, known as feature matching, represents a fundamental bottleneck in the 3D reconstruction pipeline, as it directly determines the accuracy and completeness of subsequent geometric computations. Feature matching aims to identify which features in one image correspond to the same physical point in another image, a task that becomes increasingly challenging under varying imaging conditions, geometric transformations, and environmental factors commonly encountered in UAV applications.

The complexity of feature matching stems from several inherent challenges that must be systematically addressed. Descriptors representing the same physical point may exhibit variations due to changes in viewing angle, illumination conditions, scale differences, or image noise. Additionally, repetitive patterns and ambiguous textures in natural scenes can create multiple plausible correspondences, while occlusions and dynamic elements may cause genuine features to disappear between images. UAV imagery presents particular challenges including significant scale variations due to altitude changes, rotation effects from platform movement, and atmospheric disturbances that can affect image quality and feature appearance.

2.3.1. Distance Metrics

The foundation of feature matching lies in selecting appropriate distance metrics that quantify similarity between descriptors. The choice of distance metric fundamentally determines matching quality, as different descriptor types exhibit varying sensitivities to specific distance measures based on their encoding strategies and dimensional properties.

Euclidean Distance (L2 Norm) serves as the standard metric for real-valued descriptors such as SIFT, SURF, KAZE, and DAISY, computing the square root of the sum of squared differences between corresponding descriptor elements. This metric effectively captures gradual variations in descriptor space and provides robust matching for floating-point representations that encode continuous feature properties. The L2 norm's sensitivity to outlier dimensions makes it particularly effective for descriptors that maintain consistent encoding across different viewing conditions, though computational overhead increases with descriptor dimensionality.

Hamming Distance provides the optimal metric for binary descriptors including ORB, BRISK, BRIEF, FREAK, and LATCH, measuring the number of differing bits between binary strings through efficient XOR operations. This metric's computational efficiency stems from hardware-optimized bit manipulation operations, enabling real-time matching performance crucial for UAV applications. The discrete nature of Hamming distance makes it inherently robust to small variations in binary encoding while maintaining discrimination between genuinely different features. Modern implementations leverage population count instructions (POPCNT) to achieve extremely fast distance computations, making binary descriptors particularly attractive for resource-constrained UAV platforms.

The selection between L2 and Hamming distance metrics significantly impacts both computational performance and matching accuracy, with implications extending throughout the 3D reconstruction pipeline. Floating-point descriptors paired with L2 distance typically provide superior matching accuracy under challenging conditions but require greater computational resources. Conversely, binary descriptors with Hamming distance offer exceptional computational efficiency while maintaining competitive accuracy for many UAV applications, particularly those involving real-time processing constraints.

2.3.2. Matching Strategies

Brute-Force Matching represents the most straightforward and comprehensive approach to correspondence establishment, systematically comparing every descriptor in one image with every descriptor in another image to identify the closest matches based on distance metrics. This exhaustive search strategy ensures that the optimal correspondence is identified for each feature, providing maximum accuracy when computational resources permit such comprehensive analysis.

The algorithm operates by computing distance measures between all possible descriptor pairs using the appropriate distance metric, specifically L2 for floating-point descriptors and Hamming for binary descriptors. Computational complexity represents the primary limitation of brute-force approaches, scaling quadratically with the number of features and becoming prohibitively expensive for large feature sets. However, this computational investment often yields superior matching accuracy compared to approximate methods, making brute-force matching valuable for applications where accuracy is prioritized over processing speed, including offline processing or high-precision mapping applications.

For UAV applications, brute-force matching proves particularly valuable in scenarios involving complex geometric transformations or challenging environmental conditions where approximate methods might fail to identify correct correspondences. The method's reliability makes it suitable for critical applications such as precision navigation, detailed mapping, or archaeological documentation where matching accuracy directly impacts reconstruction quality.

Fast Library for Approximate Nearest Neighbors (FLANN) represents a comprehensive framework for approximate nearest neighbor search that automatically selects optimal indexing strategies based on descriptor characteristics, dataset properties, and performance requirements (Muja & Lowe, 2009). The library provides efficient alternatives to brute-force matching through sophisticated data structures that dramatically reduce search complexity while maintaining acceptable accuracy levels for practical applications.

2.3.3. Refinement Techniques

Cross-Checking represents a powerful strategy for improving matching reliability through bidirectional correspondence verification that eliminates many false matches arising from descriptor ambiguity or algorithmic limitations. This approach performs matching operations in both directions between image pairs, retaining only correspondences that are mutually consistent across both matching directions.

The algorithm identifies potential correspondences from image A to image B, then performs reverse matching from image B to image A, accepting only feature pairs that select each other as best matches. This bidirectional verification effectively filters out many ambiguous matches that might satisfy the distance criteria in one direction but lack reciprocal support in the reverse direction.

Implementation considerations involve managing computational overhead, as cross-checking requires performing two complete matching operations for each image pair. However, this computational investment typically yields significant improvements in matching precision, reducing outlier rates and improving the reliability of subsequent geometric estimation procedures.

The effectiveness of cross-checking proves particularly valuable in challenging UAV scenarios involving repetitive textures, symmetric patterns, or marginal image quality where single-direction matching might accept ambiguous correspondences. The method provides a computationally efficient approach to outlier reduction that complements subsequent geometric verification techniques.

2.3.4. Geometric Verification

Geometric verification through robust estimation algorithms represents the final and most critical phase of feature matching, employing mathematical models to identify geometrically consistent correspondences while rejecting outliers that violate fundamental geometric constraints. This process forms the foundation for reliable 3D reconstruction by ensuring that subsequent geometric computations operate on geometrically coherent correspondence sets rather than contaminated data that could propagate errors throughout the entire reconstruction pipeline.

Random Sample Consensus (RANSAC) revolutionized robust estimation by providing a general framework for identifying consensus among data contaminated by significant outlier populations (Fischler & Bolles, 1981). The algorithm addresses the fundamental challenge of parameter estimation when the majority of observations may be corrupted by noise, mismatches, or erroneous measurements that could severely bias traditional least-squares estimation approaches.

The RANSAC paradigm operates through iterative hypothesis-and-test procedures that randomly sample minimal sets of correspondences to estimate geometric models such as fundamental matrices, homographies, or essential matrices. Each hypothesis undergoes evaluation against the complete correspondence set to identify inliers that satisfy the geometric constraints within predetermined tolerance levels. The algorithm maintains the model with the largest inlier support as the final estimate, providing robustness against outlier contamination levels that could completely overwhelm traditional estimation methods.

MAGSAC++ represents a significant advancement in robust estimation methodology, addressing several fundamental limitations of traditional RANSAC approaches through sophisticated statistical modeling and adaptive parameter selection (Baráth et al., 2020). The algorithm provides enhanced robustness through intelligent outlier handling, automatic threshold determination, and improved convergence characteristics that deliver superior performance in challenging scenarios with high outlier ratios or complex geometric relationships.

The theoretical foundation integrates marginal likelihood maximization with robust estimation principles, enabling automatic threshold selection that eliminates the need for manual parameter tuning while providing statistically principled inlier classification. This automatic threshold adaptation proves particularly valuable for UAV applications where imaging conditions vary significantly between different flight scenarios, altitudes, or environmental conditions.

Fundamental Matrix Estimation forms the cornerstone of two-view geometry, encoding the epipolar constraints that govern the geometric relationship between corresponding points across stereo image pairs (Hartley & Zisserman, 2004). The fundamental matrix provides essential geometric validation for feature correspondences while enabling subsequent camera pose estimation, 3D triangulation, and scene reconstruction applications critical for UAV-based 3D reconstuction.

Mathematical formulation establishes that corresponding points x and x' in stereo images must satisfy the epipolar constraint x'^T F x = 0, where F represents the 3×3 fundamental matrix with rank 2. This constraint provides powerful geometric validation that enables identification of correspondences consistent with the underlying camera geometry while rejecting outliers that violate geometric consistency.

For UAV applications, geometric verification proves essential for handling the significant geometric variations and potential outliers inherent in aerial imagery. The robust estimation framework provides reliable correspondence sets that serve as foundations for accurate camera pose estimation, 3D reconstruction, and navigation applications where geometric accuracy is paramount. The integration of advanced robust estimation methods with comprehensive geometric validation ensures that UAV-based reconstruction systems can handle the diverse challenging conditions encountered in operational deployments while maintaining the geometric accuracy required for practical applications.

2.4. 3D Reconstruction Theory and Pipeline Stages

Image-based 3D reconstruction represents the fundamental process through which collections of overlapping 2D images are transformed into comprehensive three-dimensional models, forming the cornerstone of UAV-based 3D reconstruction applications. This sophisticated computational framework integrates advanced computer vision principles, particularly Structure from Motion (SfM) and Multi-View Stereo (MVS) techniques, to automate the complex transformation from discrete image observations to coherent spatial representations. The methodology addresses the fundamental challenge of recovering both scene geometry and camera parameters from image data alone, enabling cost-effective and flexible approaches to detailed 3D modeling that have revolutionized applications ranging from archaeological documentation to precision agriculture.

The theoretical foundation of image-based reconstruction rests upon the geometric principle of triangulation, whereby 3D positions of scene points are determined through observations from multiple known viewpoints. However, practical UAV missions present the complex scenario where neither the precise 3D scene structure nor the exact camera positions and orientations are known a priori. The reconstruction pipeline addresses this fundamental challenge through a carefully orchestrated sequence of computational stages that simultaneously solve for both geometric unknowns while maintaining mathematical consistency and statistical robustness throughout the process.

2.4.1. Structure from Motion (SfM)

Structure from Motion constitutes the theoretical and computational foundation of modern photogrammetric reconstruction, providing elegant solutions to the classical challenges of simultaneously estimating scene geometry and camera parameters from image correspondences alone. SfM represents a paradigm shift from traditional photogrammetry by eliminating the need for pre-surveyed control points or precisely calibrated camera systems, instead employing the rich geometric constraints present in overlapping imagery to achieve reliable reconstruction through purely computational approaches.

Incremental Reconstruction Framework

The incremental approach to SfM reconstruction provides robust and controllable strategies for building comprehensive 3D models through systematic expansion from carefully selected initial configurations. This methodology addresses the fundamental challenge of initialization by identifying the most geometrically favorable image pairs based on feature correspondence quantity, spatial distribution, and geometric conditioning metrics that ensure numerical stability throughout subsequent reconstruction stages.

Initialization Phase: The reconstruction process commences with the critical selection of an optimal stereo image pair that provides strong geometric foundations for subsequent model expansion. This selection involves comprehensive analysis of correspondence characteristics, including feature distribution across image regions, baseline-to-depth ratios that optimize triangulation accuracy, and geometric conditioning measures that assess numerical stability. The relative pose estimation between the initial image pair employs robust techniques such as essential matrix decomposition, followed by careful triangulation of shared feature correspondences to establish the fundamental 3D coordinate framework.

Sequential Image Registration: Additional images are progressively integrated into the growing reconstruction through systematic pose estimation procedures that leverage the expanding set of 3D reference points. Each new image undergoes feature matching against the existing point cloud, enabling Perspective-n-Point (PnP) algorithms to estimate camera pose parameters through robust optimization techniques. This incremental approach ensures that reconstruction errors do not accumulate uncontrollably while maintaining geometric consistency across the expanding image network.

Triangulation and Point Cloud Expansion: New 3D points are systematically added to the reconstruction through triangulation of feature correspondences between newly registered images and previously processed views. This process requires careful geometric validation to ensure triangulation accuracy, including assessment of triangulation angles, reprojection error analysis, and outlier detection procedures that maintain point cloud quality while maximizing scene coverage.

Bundle Adjustment: Global Optimization Framework

Bundle adjustment represents the culmination of SfM processing, providing simultaneous refinement of all estimated parameters through comprehensive optimization that ensures global geometric consistency. This non-linear optimization procedure minimizes the cumulative reprojection error across all observations, simultaneously adjusting camera intrinsic parameters, extrinsic pose estimates, and 3D point coordinates to achieve maximum likelihood estimates under Gaussian noise assumptions.

The theoretical formulation of bundle adjustment involves constructing and solving large-scale non-linear least squares problems that capture the geometric relationships between 3D scene points and their observed 2D projections across all images. Modern implementations employ sophisticated optimization strategies including sparse matrix techniques, hierarchical solving approaches, and robust cost functions that handle outliers while maintaining computational efficiency for large-scale reconstructions involving thousands of images and millions of 3D points.

The iterative optimization process employs advanced numerical techniques such as Levenberg-Marquardt algorithms, trust region methods, and specialized sparse matrix solvers that exploit the unique structure of photogrammetric problems to achieve convergence within practical time constraints. The optimization framework incorporates sophisticated parameterization strategies for camera rotations, handling singularities and maintaining numerical stability throughout the complex parameter space exploration required for accurate reconstruction.

2.4.2. Multi-View Stereo (MVS)

While SfM provides accurate geometric frameworks through sparse feature-based reconstruction, Multi-View Stereo techniques address the critical need for comprehensive surface representation by estimating dense depth information across entire image regions. MVS algorithms leverage the precise camera pose estimates recovered through SfM to establish photometric correspondence across multiple views, enabling detailed surface reconstruction that captures fine-scale geometric detail essential for high-quality 3D modeling applications.

Dense Correspondence Estimation

MVS approaches the fundamental challenge of establishing correspondences for every pixel location within images, extending far beyond the sparse feature points utilized in SfM reconstruction. This dense correspondence estimation requires sophisticated algorithms that can reliably match pixels across multiple views while handling the complex challenges of occlusion, surface discontinuities, texture variations, and photometric inconsistencies that arise in real-world imaging scenarios.

Photometric Consistency Analysis: Dense reconstruction relies fundamentally on the assumption that corresponding surface points exhibit consistent appearance across multiple viewing directions, enabling correlation-based matching techniques to identify correct depth assignments. Advanced MVS algorithms employ normalized cross-correlation measures, mutual information metrics, and robust photometric consistency functions that account for illumination variations, surface reflectance properties, and imaging artifacts that can compromise simple intensity-based matching approaches.

Multi-View Geometric Constraints: The integration of geometric constraints from multiple viewpoints provides essential validation and refinement for dense correspondence estimates. Algorithms employ epipolar geometry constraints, multi-view triangulation consistency checks, and surface smoothness assumptions to resolve matching ambiguities and ensure geometric coherence across depth estimates from different viewing directions.

Depth Map Fusion and Surface Integration

The transformation from individual depth maps to unified surface representations requires sophisticated fusion algorithms that reconcile potentially conflicting depth estimates while preserving surface detail and maintaining global consistency. Modern approaches employ various strategies including Delaunay triangulation, implicit surface reconstruction, and volumetric fusion techniques that aggregate multi-view depth information into coherent surface models.

Confidence-Weighted Integration: Advanced fusion algorithms incorporate confidence measures that reflect the reliability of individual depth estimates, enabling intelligent combination of multi-view information while suppressing unreliable measurements. These confidence measures consider factors such as photometric consistency scores, geometric conditioning, baseline length, and surface orientation relative to viewing directions to produce weighted averages that optimize surface accuracy.

Surface Optimization and Refinement: The final surface representation undergoes comprehensive optimization procedures that enhance geometric accuracy while preserving important surface features. These refinement stages may include surface smoothing algorithms, hole-filling procedures, and mesh optimization techniques that improve surface quality while maintaining fidelity to the underlying image evidence.

2.4.3. Pipeline Integration

The integration of SfM and MVS components into cohesive reconstruction pipelines requires careful consideration of computational efficiency, memory management, and quality optimization strategies that enable practical processing of large-scale UAV datasets while maintaining reconstruction accuracy and reliability.

Computational Scalability Considerations

Modern reconstruction pipelines must address the computational challenges posed by high-resolution UAV imagery and large image collections through efficient algorithmic implementations and strategic processing approaches. Hierarchical reconstruction strategies enable processing of massive datasets through multi-scale approaches that begin with reduced-resolution imagery to establish global geometric frameworks before refining with full-resolution detail.

Memory Management and Data Structures: Efficient memory utilization becomes critical for large-scale reconstructions, requiring sophisticated data structures and caching strategies that maintain reasonable memory footprints while preserving access to essential geometric information. Modern implementations employ hierarchical data organization, streaming processing approaches, and intelligent caching mechanisms that enable processing of datasets that exceed available system memory.

Parallel Processing and Hardware Acceleration: Contemporary reconstruction systems leverage multi-core processors, GPU acceleration, and distributed computing architectures to achieve practical processing times for operational UAV applications. The inherently parallel nature of many reconstruction operations enables effective utilization of modern computational hardware through careful algorithm design and implementation optimization.

Quality Assessment and Validation Framework

Comprehensive quality assessment throughout the reconstruction pipeline ensures reliable results while enabling identification and correction of potential errors before they propagate to final models. Quality metrics encompass geometric accuracy measures, completeness assessments, and consistency validation procedures that provide quantitative evaluation of reconstruction success.

Geometric Accuracy Validation: Reconstruction accuracy assessment employs multiple validation approaches including ground control point analysis, self-consistency checks, and comparison with reference measurements where available. These validation procedures provide quantitative confidence measures that enable informed decisions about reconstruction reliability and appropriate application domains.

Completeness and Coverage Analysis: Systematic evaluation of reconstruction completeness identifies regions of insufficient coverage, poor geometric conditioning, or algorithm limitations that may compromise final model quality. This analysis guides data acquisition planning and processing parameter optimization to ensure comprehensive scene representation.

2.4.4. Contemporary Challenges

The evolution of UAV technology and expanding application requirements continue to drive research and development in 3D reconstruction methodologies, addressing emerging challenges while extending capabilities to new operational scenarios and application domains.

Challenging Environment Adaptation

Contemporary research addresses the fundamental challenges of reconstruction in complex environments that violate traditional algorithmic assumptions, including dynamic scenes with moving objects, poorly textured surfaces that limit feature detection, and extreme illumination conditions that compromise photometric consistency.

Dynamic Scene Handling: Advanced algorithms are being developed to handle dynamic elements within otherwise static scenes, employing temporal consistency constraints, motion segmentation techniques, and robust estimation procedures that can isolate and handle moving objects while preserving overall reconstruction accuracy.

Low-Texture Surface Reconstruction: Research into specialized techniques for handling poorly textured surfaces includes development of shape-from-shading approaches, integration of structured light techniques, and advanced regularization strategies that can reconstruct smooth surfaces despite limited feature correspondences.

Multi-Modal Integration and Sensor Fusion

The integration of diverse sensor modalities including LiDAR, thermal imaging, and multispectral data with traditional photogrammetric reconstruction provides enhanced capabilities and robustness through complementary information sources that can address individual sensor limitations.

RGB-LiDAR Fusion: Combined processing of optical imagery and LiDAR data enables reconstruction approaches that leverage the geometric accuracy of laser scanning with the texture detail and completeness of photogrammetric reconstruction, producing hybrid models that exceed the capabilities of either approach individually.

Thermal and Multispectral Integration: Integration of thermal and multispectral imagery with traditional reconstruction provides enhanced surface characterization capabilities while enabling reconstruction in challenging visual conditions where traditional approaches might fail due to poor illumination or limited texture contrast.

The comprehensive understanding of 3D reconstruction theory and pipeline stages provides essential foundation for evaluating and optimizing the feature extraction and matching algorithms that form the critical initial stages of the reconstruction process. The quality and reliability of feature detection, description, and matching directly impact all subsequent reconstruction stages, making objective algorithm evaluation essential for achieving optimal reconstruction performance across diverse UAV applications and operational scenarios.

2.5. Evaluation Methodologies

The evaluation of feature extraction and matching algorithms has historically relied on subjective criteria and limited performance metrics, creating significant gaps in objective assessment methodologies. Traditional evaluation frameworks often focus on individual performance aspects such as repeatability, distinctiveness, or computational requirements, but fail to provide frameworks that integrate multiple criteria into unified assessment scores suitable for practical deployment decisions.

2.5.1. Limitations of Existing Evaluation Approaches

Most existing evaluation frameworks suffer from several limitations that have become increasingly apparent in the contemporary research landscape. First, they typically evaluate algorithms on single datasets or limited environmental conditions, providing insufficient evidence for generalization across diverse scenarios encountered in real-world UAV applications.

Second, traditional evaluation methodologies rely on manually weighted combinations of performance metrics, introducing subjective bias that undermines objective comparison. This subjective weighting becomes problematic when attempting to establish fair comparisons between fundamentally different algorithmic approaches.

Third, existing frameworks rarely address cross-platform performance variations, despite the increasing importance of deployment across diverse computational environments ranging from high-performance servers to resource-constrained UAV platforms.

2.5.2. Contemporary Evaluation Challenges

The emergence of deep learning-based feature extraction and matching methods has introduced additional evaluation complexities that existing frameworks inadequately address. Deep learning methods often require different performance metrics, evaluation protocols, and computational considerations compared to traditional approaches. The lack of standardized evaluation frameworks creates challenges for practitioners attempting to make informed algorithm selection decisions.

Furthermore, the increasing emphasis on end-to-end learning approaches that bypass traditional feature extraction stages entirely requires new evaluation methodologies that can assess performance across different levels of algorithmic abstraction. Traditional evaluation metrics focused on keypoint repeatability and descriptor distinctiveness may not adequately capture the performance characteristics of methods that establish correspondences directly without explicit feature extraction.

Research Gap: Objective Algorithm Assessment

The development of objective, comprehensive evaluation methodologies represents a research gap that transcends the specific choice between traditional and deep learning approaches. This gap is particularly pronounced in the context of UAV applications, where algorithm selection must consider multiple competing performance objectives including accuracy, computational efficiency, and deployment complexity.

Current evaluation practices often rely on ad-hoc combinations of performance metrics without rigorous statistical validation or consideration of metric interdependencies. This approach fails to provide the systematic, evidence-based guidance required for practical algorithm selection in operational UAV systems where multiple constraints must be simultaneously satisfied.

2.5.3. Novel Contribution: Composite Unsupervised Efficiency Score

This research addresses these gaps through the development of the novel CUES methodology. The methodology provides statistically validated, bias-free algorithm evaluation that supports evidence-based selection for practical UAV applications.

The proposed framework offers several key advantages over existing evaluation approaches. First, it eliminates subjective bias through the integration of multiple unsupervised weighting methodologies that objectively determine the relative importance of different performance criteria. Second, it provides assessment across diverse datasets and environmental conditions, ensuring reliable generalization. Third, it explicitly addresses cross-platform performance variations, enabling informed deployment decisions across different computational environments.

Beyond the core scoring methodology, CUES incorporates rigorous statistical validation components that address limitations in existing evaluation frameworks. Bootstrap-based 95% confidence interval estimation enables uncertainty quantification for individual algorithm combinations, providing statistical significance testing between competing alternatives rather than relying on point estimates alone. Additionally, comprehensive parameter sensitivity analysis validates the robustness of methodological choices, demonstrating that algorithmic rankings remain stable across variations in normalization parameters and weight aggregation strategies. These validation components ensure that conclusions derived from the evaluation framework reflect genuine algorithmic characteristics rather than artifacts of specific parameter selections.

2.5.4. Academic and Practical Relevance

The evaluation framework developed in this research maintains relevance regardless of the underlying algorithmic paradigm, whether traditional hand-crafted methods or emerging deep learning approaches. The methodology focuses on efficiency outcomes rather than algorithmic implementation details, enabling fair comparison across different methodological approaches while providing practical guidance for algorithm selection in resource-constrained UAV applications.

This comprehensive evaluation approach addresses a gap in the literature by providing an objective, statistically validated methodology for multi-criteria algorithm assessment that can inform both academic research and practical deployment decisions. The framework's emphasis on cross-platform evaluation and multi-criteria optimization addresses the complex requirements of modern UAV applications while maintaining applicability to future algorithmic developments.

2.5.5. Unsupervised Weighting Methods

The development of unsupervised weighting methods represents an advancement in objective MCDM applications, providing data-driven approaches for criterion importance assessment that eliminate subjective bias while maintaining theoretical rigor. These methods derive weighting schemes directly from performance data characteristics, ensuring that evaluation outcomes reflect empirical evidence rather than subjective preferences.

Entropy Weighting Based on Information Theory

Entropy weighting methodology derives from Shannon's information theory, measuring the discriminative power of evaluation criteria by quantifying the information content provided by each criterion for distinguishing between alternatives (Li et al., 2011; Shannon, 1948). The approach assigns higher weights to criteria exhibiting lower entropy values, indicating greater information content and improved discriminative capability across the evaluation space.

The theoretical foundation recognizes that criteria with uniform distributions across alternatives provide limited discriminative information, while criteria showing significant variation effectively differentiate between alternative performances. Mathematical implementation involves probability distribution calculation for each criterion, entropy computation using Shannon's formula, and weight derivation through inverse entropy scaling that emphasizes discriminative criteria.

Applications in algorithm evaluation demonstrate that entropy weighting effectively identifies efficiency metrics that provide meaningful differentiation between algorithmic approaches, supporting objective algorithm selection while maintaining theoretical consistency with information-theoretic principles. The method's emphasis on discriminative power makes it particularly valuable for identifying criteria that capture genuine efficiency differences rather than random variation.

CRITIC Method: Contrast and Correlation Analysis

The Criteria Importance Through Intercriteria Correlation (CRITIC) method provides an approach that combines two fundamental aspects of criterion importance: contrast intensity through standard deviation analysis and conflict assessment through intercriteria correlation evaluation (Diakoulaki et al., 1995). This dual consideration ensures that selected criteria provide both discriminative power and independent information content.

Theoretical foundations recognize that effective evaluation criteria should exhibit high variability across alternatives while maintaining low correlation with other criteria, ensuring that each criterion contributes unique information to the evaluation process. The CRITIC methodology quantifies these properties through statistical analysis, computing contrast intensities and correlation matrices to identify criteria that maximize both discrimination and independence.

Principal Component Analysis Weighting

PCA weighting methodology utilizes dimensionality reduction principles to identify evaluation criteria that contribute most significantly to the overall variance structure of efficiency data (Wold et al., 1987). The approach analyzes underlying patterns in algorithm efficiency to emphasize criteria that capture the most important dimensions of variation across the evaluation space.

Theoretical foundations derive from multivariate statistics and the recognition that algorithm efficiency often exhibits underlying structural patterns that can be captured through variance analysis. PCA weighting identifies these patterns and emphasizes criteria that align with the principal components representing maximum variance, ensuring that evaluation outcomes reflect the most significant sources of efficiency variation.

Variance Weighting Methodology

Variance weighting represents the most direct approach to discriminative power assessment, assigning criterion importance weights proportional to the variance exhibited across alternative performances (Odu, 2019). This straightforward methodology emphasizes criteria that effectively differentiate between alternatives while de-emphasizing criteria showing limited variation.

The theoretical foundation recognizes that evaluation criteria providing limited discrimination contribute minimal value to algorithm selection decisions, while criteria exhibiting significant variance enable meaningful differentiation between algorithmic approaches. Variance weighting directly quantifies this discriminative capability through statistical variance analysis.

Research demonstrates variance weighting's effectiveness as both a standalone method and as a component in composite weighting frameworks, where its straightforward discriminative emphasis complements more sophisticated approaches that consider additional factors such as correlation and structural patterns.

This chapter has established the theoretical foundation and identified the critical evaluation gaps that motivate the development of the comprehensive evaluation framework presented in this research. The review demonstrates that while significant advances have been made in both traditional and deep learning-based feature extraction methods, the fundamental challenge of objective, comprehensive evaluation for practical deployment scenarios remains largely unaddressed across all algorithmic paradigms.

3. MATERIAL AND METHOD

This chapter details the materials and methods employed in evaluating feature detection, description, and matching algorithms for UAV-based 3D reconstruction. The research methodology utilizes five datasets including synthetic transformations, standard benchmarks, real-world UAV imagery, and simulation-generated content across diverse operational conditions.

3.1. Materials

3.1.1. Hardware Platforms

The research employed dual computational platforms to assess algorithm performance across different hardware architectures representing typical UAV deployment scenarios. This approach evaluates computational efficiency patterns and ranking consistency between server-grade and mobile computing environments.

Primary Platform: Intel Xeon Gold 6130 CPU @ 2.10GHz (60-core server), 100 GB DDR4 memory, Ubuntu 24.04.2 LTS. This configuration represents ground station processing environments and dedicated photogrammetric workstations.

Secondary Platform: ARM-based mobile processors including OnePlus 7 (Snapdragon 855) and Xiaomi Redmi Note 13 Pro+ (Mediatek Dimensity 7200 Ultra) representative of edge computing platforms and resource-constrained environments relevant for real-time UAV applications and field deployment.

Both platforms operated identical software environments (OpenCV 4.11, Python 3.8) to ensure fair comparisons and eliminate software-induced performance variations.

3.1.2. Software Framework

The research implementation utilizes Python 3.8 as the primary development environment, providing comprehensive access to computer vision libraries, numerical computing frameworks, and statistical analysis tools required for extensive algorithm evaluation. OpenCV 4.11 serves as the core computer vision library, offering standardized implementations of feature detection, description, and matching algorithms that ensure reproducible results across different research environments.

COLMAP integration through pycolmap 3.11.1 enables programmatic 3D reconstruction pipeline execution with consistent parameter configurations across all algorithm combinations. This integration provides access to state-of-the-art Structure from Motion algorithms while maintaining automated processing capabilities required for systematic evaluation of 784 distinct algorithm combinations.

Microsoft AirSim 1.8.1, built on Unreal Engine, provides the simulation environment for generating synthetic UAV datasets with precise control over flight parameters, camera configurations, and environmental conditions. The simulation framework enables systematic dataset generation with ground truth information while maintaining realistic visual characteristics representative of UAV imagery.

Data processing and analysis utilize NumPy for numerical computations, Pandas for data manipulation, and Plotly for interactive visualization generation. Statistical analysis employs SciPy for correlation analysis and hypothesis testing required for methodology validation. The complete software framework ensures reproducible research outcomes while supporting comprehensive statistical evaluation of algorithm performance patterns.

3.1.3. Datasets

Five distinct datasets provide evaluation coverage across diverse UAV-based 3D reconstruction scenarios:

Synthetic Dataset: Generated from controlled image transformations, evaluating algorithm robustness to rotation, scaling, and intensity changes with exact ground truth information. The source image was selected from the HPatches dataset (Balntas et al., 2017), specifically the "woman" sequence (v_woman/1.jpg), showing a scene containing a challenging mix of textures like rough brick and smooth paint, vibrant colors, and complex structures including sharp edges and overlapping forms Figure 3.1.

Oxford Affine Dataset: The Oxford Affine Covariant Features dataset (Mikolajczyk et al., 2005) is a widely used benchmark for evaluating local feature detectors and descriptors. It consists of image sequences depicting various scenes subject to significant viewpoint changes, scale variations, and illumination changes. The dataset encompasses five different changes in imaging conditions: viewpoint changes, scale changes, image blur, JPEG compression, and illumination variations Figure 3.2. In viewpoint change tests, the camera varies from a fronto-parallel view to one with significant foreshortening at approximately 60 degrees to the camera. Scale change and blur sequences are acquired by varying camera zoom and focus respectively, with scale changes by approximately a factor of four. Each test sequence contains 6 images with gradual geometric or photometric transformation at medium resolution 800 × 640 pixels. This established benchmark provides standardized evaluation scenarios with significant viewpoint changes, scale variations, and illumination modifications for robust comparison with existing research.

UAV Dataset: Curated real-world imagery from 10 Pix4D drone projects representing diverse UAV operational scenarios with varying environmental conditions and scene complexity Figure 3.3. Two consecutive images were manually selected from each project to form sequential evaluation pairs. Images were resized to 1000 pixels wide while maintaining aspect ratio to standardize processing across varying camera resolutions.

AirSim Dataset: Controlled simulation environment using Microsoft AirSim enabling systematic evaluation under controlled flight parameters while maintaining realistic visual characteristics Figure 3.4.

Figure 3.1 Selected image for Synthetic Dataset creation: woman

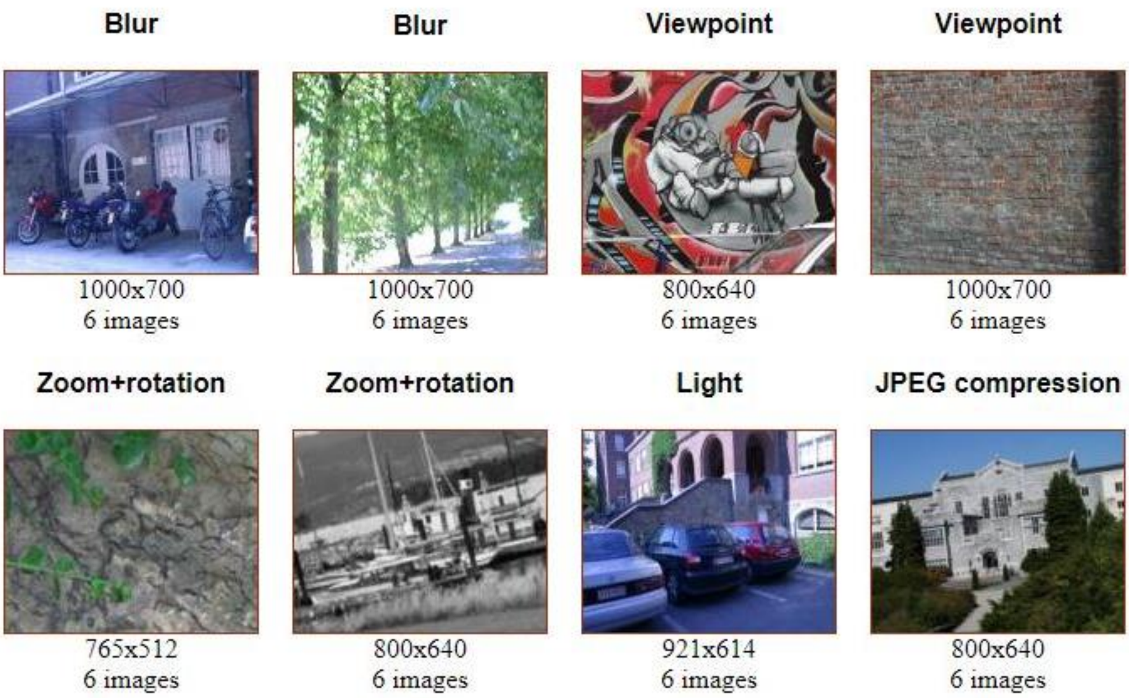


Figure 3.2 Oxford Dataset 8 scenarios

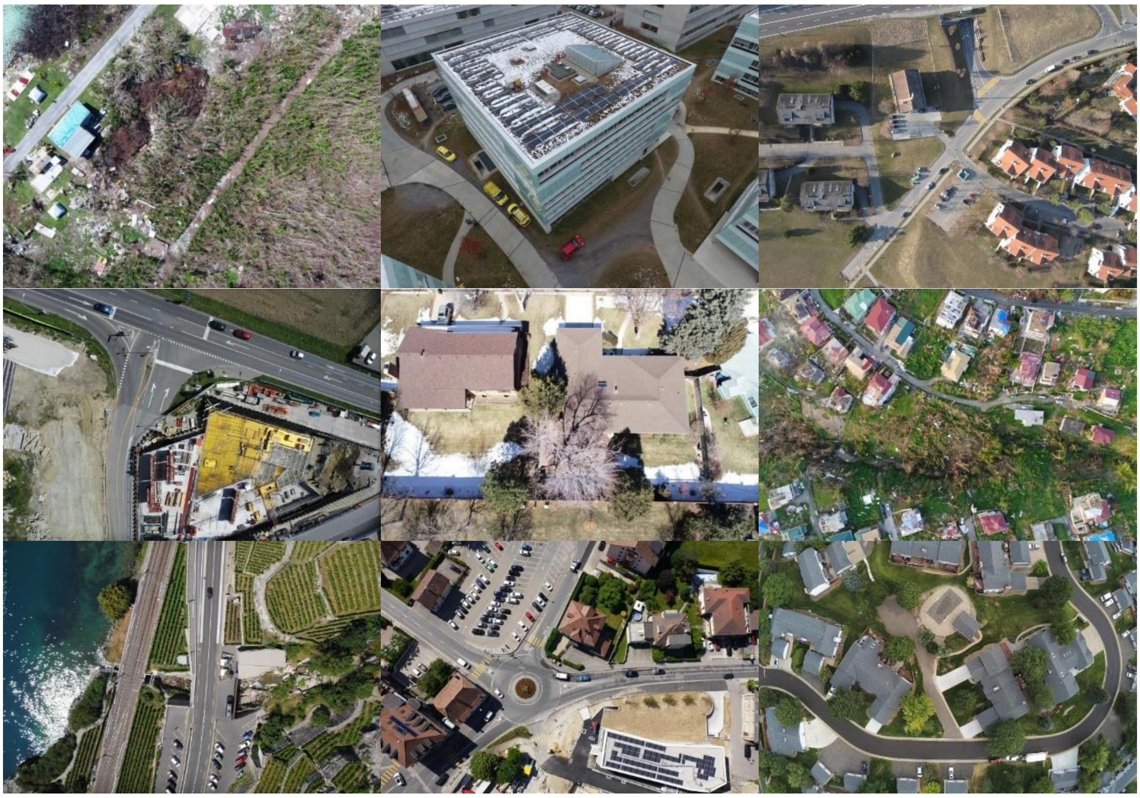


Figure 3.3 UAV Dataset - Selected Pix4D projects

Figure 3.4 AirSimNH Environment

Drone Dataset: Complete UAV flight imagery from a single Pix4D building project providing comprehensive 3D reconstruction evaluation under realistic operational conditions. This dataset represents a realistic UAV photogrammetry scenario, capturing images of a small building from multiple viewpoints with image overlap ranging from 50% to 80%, representative of typical UAV data acquisition practices for 3D modeling Figure 3.5.

Figure 3.5 Drone Dataset

3.2. Methods

3.2.1. Synthetic Dataset Generation Protocol

The synthetic dataset generation protocol employs systematic image transformation procedures to evaluate algorithm robustness under controlled conditions. A single source image selected from the HPatches dataset provides the foundation for generating transformation sequences that isolate specific sources of variation commonly encountered in UAV imagery processing.

As described in Section 3.1.3, the source image selection criteria emphasize challenging visual characteristics including complex textures, varied illumination, sharp edges, and overlapping geometric structures that represent typical UAV scene content. The selected woman image from the HPatches dataset contains a mixture of building materials, vegetation, and architectural features that provide comprehensive feature detection challenges across different algorithm approaches.

Geometric transformations include rotation sequences applied at 15-degree increments from 15 to 90 degrees, scale variations ranging from 0.5 to 1.5 with intermediate factors of 0.7, 0.9, 1.1, and 1.3, providing systematic evaluation of scale invariance properties. Each transformation maintains image quality while preserving ground truth correspondences that enable quantitative accuracy assessment.

Photometric transformations encompass both additive and multiplicative intensity variations designed to simulate illumination changes encountered in UAV operations. Additive changes apply constant offsets of ±10 and ±30 pixel values, while multiplicative changes employ scaling factors of 0.7, 0.9, 1.1, and 1.3 to simulate exposure variations and atmospheric effects.

The controlled transformation approach enables precise evaluation of algorithm responses to individual variation sources while maintaining exact ground truth information for quantitative accuracy assessment. This methodology provides fundamental insights into algorithm robustness characteristics that inform practical deployment decisions across diverse operational conditions.

3.2.2. AirSim Dataset Acquisition Protocol

The AirSim dataset acquisition employs Microsoft AirSim simulation environment to generate comprehensive synthetic UAV imagery with precise control over flight parameters, camera configurations, and environmental conditions. The simulation approach enables systematic evaluation under controlled conditions while maintaining realistic visual characteristics representative of UAV operations.

The simulation environment utilizes the default Neighborhood setting within AirSim, providing a realistic suburban landscape with residential buildings, road networks, vegetation, and varied terrain features representative of typical UAV mapping scenarios. This environment offers sufficient complexity for algorithm evaluation while maintaining computational efficiency for large-scale dataset generation.

Two AirSim datasets are generated and released publicly, targeting different evaluation needs.

AirSimNH dataset (multi-view, grid survey): This dataset is created using the airsim_collect.py script. Image acquisition is implemented as a grid-based photogrammetry survey with multiple altitude layers and multiple camera orientations per grid node. The dataset is organized into three altitude layers (AirSim NED coordinates) stored as Layer1 (z=-60), Layer2 (z=-100), and Layer3 (z=-140), corresponding to approximately 60 m, 100 m, and 140 m altitude in the simulator.

Within each layer, the UAV visits evenly spaced (x, y) grid nodes with a 40-unit step size, producing different coverage densities: 9x9 nodes for Layer1 (range -160 to 160), 7x7 nodes for Layer2 (range -120 to 120), and 5x5 nodes for Layer3 (range -80 to 80). At each grid node, five fixed camera view configurations are captured to provide viewpoint diversity, stored as separate folders with pitch/yaw pairs: (-60, 0), (-60, -90), (-90, 0), (-120, 0), and (-120, -90). The pitch -60 views represent the natural forward-looking drone camera angle, while pitch -90 represents the nadir view used for mapping-style capture. Images are saved as PNG files with filenames encoding the simulated pose in the form id_zZ_yY_xX.png.

```
Input:
  - AirSim simulation environment (connected via VehicleClient)
  - Output directory path (tmp_dir)
  - Photogrammetry parameters: Field of View (FOV), Step size (STEP), Overlap (OVERLAP)
  - Layer configurations: altitude, grid size, and coordinate points for each layer
  - Camera pitch angle (PITCH_ANGLE)
Output:
  - High-resolution images saved to specified subdirectories for each layer
Initialize AirSim client connection.
Define and create the output directory for storing captured images.
Define photogrammetry constants: FOV, STEP, OVERLAP.
Define layer_configs dictionary:
  - Each key (1, 2, 3) represents a layer.
  - Each layer contains:
    - size: Grid dimension (e.g., 3 for 3x3, 5 for 5x5, 7 for 7x7).
    - altitude: Negative Z-coordinate for flight height.
    - points: List of X/Y coordinates defining grid nodes for that layer.
Define global camera PITCH_ANGLE (e.g., -60 degrees for forward-looking oblique).
Initialize image_counter to 1.
for each layer (from 1 to 3) do
    Retrieve config (altitude, size, points) for the current layer.
    Calculate ground_coverage (effective ground area covered by an image at current altitude).
    Calculate effective_overlap based on STEP and ground_coverage.
    Print current layer configuration details (altitude, ground coverage, effective overlap).
    Initialize moving_right boolean to TRUE (for serpentine pattern).
    for each y_coordinate in config.points do
        if moving_right is TRUE then
            Set x_range to config.points (traverse from left to right).
            Set YAW_ANGLE to 0 degrees (forward).
        else (moving_right is FALSE)
            Set x_range to config.points in reverse order (traverse from right to left).
            Set YAW_ANGLE to 180 degrees (backward).
        end if
        for each x_coordinate in x_range do
            Construct pose (position and orientation) for the drone:
              - Position: (x_coordinate, y_coordinate, altitude).
              - Orientation: Quaternion derived from PITCH_ANGLE, 0 roll, and YAW_ANGLE.
            Set drone's pose in the AirSim simulation to the calculated pose.
            Pause briefly (e.g., 0.1 seconds) to allow stabilization.
            Request a "Scene" image from the "front_center" camera.
            Create a layer-specific subdirectory within the tmp_dir.
            Construct image_filename using image_counter, altitude, y_coordinate, x_coordinate.
            Save the captured image data to image_filename within the layer subdirectory.
            Increment image_counter.
            Print confirmation of saved image.
        end for
        Toggle moving_right (reverse direction for next row).
    end for
end for
Reset drone's pose to origin (0, 0, 0) with default orientation in AirSim.
```

Figure 3.6 AirSimNH-natural Single-View Serpentine Acquisition Pseudocode

AirSimNH-natural dataset (single-view, mission-like traverse): This dataset is created using the airsim_collect_natural.py script and focuses on a natural drone view with a fixed pitch angle of -60 degrees. Only a single camera view is captured at each grid node. The acquisition follows a mission-like serpentine traversal, where yaw alternates between 0 and 180 degrees to maintain a forward-facing camera direction along the flight path. Three layers are generated with decreasing grid density (7x7, 5x5, and 3x3), using simulated altitudes around z=-80, z=-110, and z=-140, respectively.

Dataset used in evaluation: For the feature-level evaluation reported in this thesis, only a reduced subset of AirSimNH is used to limit computation. Specifically, the benchmark uses the nadir subset Layer3/pitch-90_yaw0 (25 images arranged as a 5x5 grid at approximately 140 m height). Matching is performed between consecutive images in the sorted filename order (image n with n+1), consistent with the implemented evaluation script.

Camera configuration maintains consistent parameters including 1000 x 665 pixel resolution, 90-degree field of view, and automatic exposure settings with motion blur disabled. These settings provide realistic image characteristics while ensuring consistent evaluation conditions across all generated datasets.

The simulation framework provides ground truth camera poses and known acquisition geometry through the recorded (x, y, z) positions. In this thesis, the feature-level evaluation relies on image-based geometric verification (fundamental matrix inliers) rather than direct use of simulator ground truth, while still benefiting from the controlled and repeatable acquisition conditions.

3.2.3. Algorithm Implementation Protocol

The algorithm implementation protocol ensures systematic evaluation of all feasible combinations of feature detection, description, and matching techniques available in OpenCV 4.11. This comprehensive approach encompasses 784 distinct algorithm combinations across multiple evaluation scenarios to provide complete coverage of algorithmic possibilities within the OpenCV framework. The core evaluation loop structure is illustrated in Figure 3.7.

Feature detector selection includes FAST, AGAST, SIFT, ORB, BRISK, MSER, KAZE, AKAZE, GFTT, GFTT with Harris enhancement, STAR, Harris-Laplace, MSD, and TBMR algorithms. Each detector implementation utilizes optimized parameters determined through systematic parameter space exploration to ensure fair comparison and maximum performance potential for each algorithm.

Feature descriptor pairing encompasses SIFT, ORB, BRISK, FREAK, BRIEF, AKAZE, DAISY, LUCID, LATCH, VGG, BEBLID, TEBLID, and BOOST descriptors. Compatibility constraints between binary and floating-point descriptors are properly handled through appropriate distance metric selection and matching strategy configuration. The inlier evaluation procedure utilizing geometric verification is presented in Figure 3.8.

Matching strategy evaluation includes both Brute-Force and FLANN-based approaches with appropriate distance metrics including L2 norm for floating-point descriptors and Hamming distance for binary descriptors. Cross-checking and ratio testing parameters are optimized for each matching approach to ensure robust correspondence establishment.

Geometric verification employs MAGSAC++ algorithm through OpenCV's findFundamentalMat function to eliminate outlier correspondences and identify geometrically consistent matches. This approach provides reliable fundamental matrix estimation while maintaining consistency across all algorithm combinations and datasets.

```
image_folder ← "./Datasets/AirSimNH/Layer3/pitch-90_yaw0"
images ← Load all images from image_folder.
Initialize or load PerformanceMetrics array (Rate).
Initialize or load ExecutionTimes array (Exec_time).
for k from 0 to len(images)-2 do                    ▷ Iterate through consecutive image pairs
    img1 ← images[k]
    img2 ← images[k+1]
    for i from 0 to len(Detectors)-1 do             ▷ Iterate through each detector type
        current_detector ← Detectors[i].
        keypoints1 ← current_detector.detect(img1).
        Record start_time for detection.
        keypoints2 ← current_detector.detect(img2).
        detect_time ← Elapsed time for detection.
        for j from 0 to len(Descriptors)-1 do       ▷ Iterate through each descriptor type
            current_descriptor ← Descriptors[j].
            for norm_type_idx from 0 to 1 do        ▷ Norm. Type (0: L2, 1: Hamming)
                for match_type_idx from 0 to 1 do   ▷ Matching (0: BF, 1: FLANN)
                    normalization_method ← Normalization[norm_type_idx].
                    descriptors1 ← current_descriptor.compute(img1, keypoints1).
                    Record start_time for description.
                    descriptors2 ← current_descriptor.compute(img2, keypoints2).
                    descript_time ← Elapsed time for description.
                    Record start_time for matching and outlier rejection.
                    inliers, matches ← evaluate_with_fundamentalMat_and_XSAC().
                    match_and_ransac_time ← Elapsed time for matching.
                    Store match_and_ransac_time in Exec_time.
                    Rate, Exec_time ← process_matches().    ▷ Calculates metrics
                    Mark relevant entries in Exec_time and Rate as invalid.
                end for
            end for
        end for
    end for
end for
```

Figure 3.7 Core Feature Evaluation Loop Pseudocode

Caching mechanisms optimize computational efficiency by storing detected keypoints and computed descriptors across multiple matching operations. This approach eliminates redundant computations while ensuring consistent timing measurements that accurately reflect practical deployment scenarios. Algorithm combinations that fail at any stage like no detected keypoints, descriptor computation errors, empty descriptors, or matching failures are recorded as invalid for that image pair and excluded from dataset-level averages, enabling the reporting of valid evaluation counts in the Results chapter.

```
function evaluate_with_fundamentalMat_and_XSAC
    if matcher_type is Brute-Force (0) then
        bf_matcher ← Create Brute-Force Matcher and cross-check.
        matches ← bf_matcher.match(Descriptors1, Descriptors2).
    else (matcher_type is FLANN-Based (1))
        if normalization_type is L2 then
            index_params ← K-D Tree, trees=5.
        else (normalization_type is Hamming)
            index_params ← LSH, table_number=6, key_size=12, multi_probe_level=1.
        end if
        search_params ← Dictionary with checks=50.
        flann_matcher ← Create FLANN-Based Matcher.
        matches1to2 ← flann_matcher.match(Descriptors1, Descriptors2).
        matches2to1 ← flann_matcher.match(Descriptors2, Descriptors1).
        matches ← Perform cross-check on matches1to2 and matches2to1.
    end if
    points1 ← Extract 2D coordinates from Keypoints1 based on matches.
    points2 ← Extract 2D coordinates from Keypoints2 based on matches.
    mask ← Compute F Matrix from points1 and points2 using cv2.USAC_MAGSAC.
    inliers ← Filter matches where corresponding mask value is 1 (indicating inlier).
    return inliers, matches.
end function
```

Figure 3.8 Inlier evaluation function Pseudocode with Fundamental Matrix and MAGSAC++

3.2.4. COLMAP Integration Protocol

The COLMAP integration protocol enables systematic 3D reconstruction evaluation for each algorithm combination through programmatic database generation and incremental reconstruction execution. This approach provides quantitative assessment of reconstruction quality while maintaining consistent processing parameters across all evaluated combinations.

Database generation creates dedicated COLMAP databases for each unique detector-descriptor-matcher combination using pycolmap 3.11.1 interface. Feature information and match correspondences are programmatically inserted into each database following COLMAP's data structure requirements while maintaining compatibility with existing COLMAP workflows.

Incremental reconstruction employs COLMAP's established SfM pipeline through pycolmap.incremental_mapping() function with consistent initialization parameters across all combinations. Initial image pair selection utilizes predetermined image indices to ensure consistent reconstruction starting conditions while bundle adjustment parameters maintain default COLMAP settings optimized for photogrammetric applications.

Reconstruction metrics include mean reprojection error computed through COLMAP's internal functions, providing quantitative assessment of geometric accuracy achieved by each algorithm combination. Point cloud density metrics quantify reconstruction completeness through 3D point counts successfully triangulated during the reconstruction process.

The automated pipeline handles reconstruction failures gracefully while maintaining comprehensive logging of processing outcomes and error conditions. This approach enables systematic evaluation of algorithm reliability while providing insights into practical deployment considerations for operational environments.

Quality assessment encompasses both geometric accuracy through reprojection error analysis and reconstruction completeness through point density evaluation. These metrics provide comprehensive insights into practical reconstruction performance while supporting algorithm selection decisions for specific application requirements.

3.2.5. Timing Calculation Methodology

The timing calculation methodology provides comprehensive assessment of computational efficiency across different algorithm components and operational scenarios. Precise timing measurements enable identification of computational bottlenecks while supporting algorithm selection decisions for resource-constrained deployment environments.

Raw processing time measurements encompass detection time for keypoint identification, description time for feature vector computation, and matching time for correspondence establishment. Total processing time combines all pipeline components to provide comprehensive assessment of complete feature processing workflows.

Normalized timing metrics address variations in feature counts across different algorithm combinations and image content. The 1k Total Time metric normalizes total processing time per 1000 detected features, providing fair comparison across algorithms with different feature density characteristics. The 1k Inlier Time metric normalizes processing time per 1000 geometrically verified correspondences, reflecting computational efficiency for producing reliable matches suitable for 3D reconstruction.

Cross-platform timing evaluation assesses computational efficiency consistency between server-grade and mobile computing environments. Timing measurements utilize identical algorithm implementations and parameter configurations across both platforms to ensure fair comparison while revealing platform-specific optimization opportunities.

Statistical analysis of timing measurements includes confidence interval estimation and outlier detection to ensure robust efficiency assessment. Multiple timing runs provide statistical reliability while systematic timing protocols minimize measurement variance across different evaluation conditions.

The timing methodology supports identification of optimal algorithm combinations for different deployment scenarios including real-time mobile applications requiring minimal computational overhead and high-precision server applications where accuracy takes precedence over processing speed.

3.2.6. Composite Efficiency Score Methodology

The composite efficiency score methodology integrates multiple unsupervised weighting approaches to provide objective, comprehensive algorithm evaluation without subjective bias in criterion importance assignment. This approach combines entropy, CRITIC, PCA, and variance weighting methods through principled mathematical combination that maintains theoretical consistency while eliminating researcher-imposed preferences.

Entropy weighting quantifies information content provided by each evaluation criterion through Shannon's information theory framework. Criteria exhibiting greater variation across algorithm alternatives receive higher weights, emphasizing performance dimensions that distinguish between algorithmic approaches. Mathematical implementation computes probability distributions for each criterion, calculates entropy values using Shannon's formula, and derives weights through inverse entropy scaling.

CRITIC weighting integrates contrast intensity through standard deviation analysis with conflict assessment through intercriteria correlation evaluation. This dual approach ensures that selected criteria provide both discriminative power and independent information content. Implementation involves standard deviation calculation for contrast assessment, correlation matrix computation for independence evaluation, and composite scoring that emphasizes criteria exhibiting high contrast and low correlation.

PCA weighting uses dimensionality reduction principles to identify evaluation criteria contributing most significantly to overall variance structure. The approach analyzes underlying patterns in algorithm performance to emphasize criteria capturing important dimensions of variation. Mathematical implementation involves covariance matrix computation, eigenvalue decomposition for principal component identification, and weight derivation based on component loadings.

Variance weighting provides direct assessment of discriminative power by assigning criterion importance weights proportional to variance exhibited across algorithm performances. Implementation simplicity represents a key advantage while maintaining effectiveness for algorithm evaluation scenarios where discriminative power represents the primary consideration.

Composite efficiency score calculation combines individual weighting methods through arithmetic averaging to balance different theoretical perspectives while maintaining mathematical consistency. This approach prevents dominance by any single weighting method while preserving the discriminative advantages of each individual approach.

The nonlinear normalization function employs a square root transformation with alpha=0.5 following the formula x_norm = ((x - min) / (max - min))^alpha. This parameter choice is grounded in multi-criteria decision making literature and provides balanced sensitivity across the performance spectrum. The square root transformation reduces the impact of extreme outliers while preserving rank order. More linear normalization (alpha closer to 1) would preserve original performance relationships but may compress important distinctions between algorithms, while more aggressive transformations with alpha<0.5 risk over-emphasizing minor differences. Sensitivity analysis across alpha values from 0.1 to 0.75 confirms that CUES rankings demonstrate exceptional stability with Spearman rho averaging 0.9968 across all datasets, validating alpha=0.5 as a robust default that produces consistent algorithmic hierarchies regardless of parameter variations.

The arithmetic mean aggregation of entropy, CRITIC, PCA, and variance weights follows the formula w_combined = (w_entropy + w_pca + w_critic + w_variance) / 4. This choice reflects equal treatment of all weighting methods without prior preference, ensuring transparency and ease of interpretation. Each component method captures distinct aspects of the data structure: entropy quantifies information content through discriminative power, PCA identifies variance-maximizing dimensions, CRITIC measures correlation conflict for independence assessment, and variance directly measures distribution breadth. Comparative analysis with geometric mean, harmonic mean, and trimmed mean aggregation methods demonstrates near-perfect rank correlation with Spearman rho averaging 0.9996 across all datasets, confirming that final algorithm rankings remain essentially invariant to the aggregation method choice. The arithmetic mean is therefore preferred for its computational simplicity, intuitive interpretation, and demonstrated equivalence to alternative approaches.

3.2.7. Evaluation Metrics Framework

The evaluation metrics framework encompasses comprehensive assessment dimensions including matching quality, computational efficiency, and reconstruction accuracy. This multi-dimensional approach provides complete characterization of algorithm performance while supporting evidence-based selection decisions for diverse deployment scenarios.

Matching quality metrics assess the performance and reliability of correspondence establishment between image pairs. Precision quantifies the proportion of established matches that are geometrically correct, calculated as the ratio of verified inliers to total putative matches. Recall measures the completeness of producing usable correspondences relative to detected features in the reference image, computed as the ratio of verified inliers to the number of detected keypoints in the first image of the evaluated pair.

Repeatability evaluates consistency of feature detection across different viewpoints or transformations, quantified as the ratio of verified inliers to the minimum number of detected keypoints between the two images. F1-Score provides balanced assessment through harmonic mean of precision and recall, offering single metric characterization of matching performance.

Computational performance metrics quantify processing speed across different pipeline components and normalization approaches. Raw timing measurements include detection time, description time, matching time, and total processing time for comprehensive workflow assessment. Normalized metrics including 1k Total Time and 1k Inlier Time provide fair comparison across algorithms with different feature density characteristics.

Reconstruction accuracy metrics assess geometric quality of 3D models generated through SfM pipelines. Mean reprojection error quantifies geometric consistency between 3D point projections and observed image coordinates. Point cloud density metrics quantify reconstruction completeness through successfully triangulated 3D point counts.

The comprehensive metrics framework enables multi-dimensional algorithm characterization while supporting objective comparison across diverse evaluation criteria. Statistical analysis of metrics distributions provides confidence assessment while correlation analysis reveals relationships between different performance dimensions.

4. EXPERIMENTAL DESIGN

This chapter presents the evaluation framework for 784 algorithm combinations across five datasets with dual-platform testing. The design addresses fair algorithm comparison through standardized parameter optimization, consistent metrics, and statistical validation. Results support evidence-based algorithm selection for UAV applications.

4.1. Algorithm Evaluation Matrix

4.1.1. Detector Selection

Fourteen feature detectors from OpenCV 4.11 undergo evaluation, including corner detection, scale-space, and binary methods. Each detector receives optimized parameters to eliminate selection bias. Table 4.2 presents the OpenCV parameters for methods capable of both detection and description, while Table 4.1 lists parameters for detection-only methods.

Table 4.1 Detectors only methods' OpenCV parameters

FAST	cv2.FastFeatureDetector_create(nonmaxSuppression=True, type=cv2.FAST_FEATURE_DETECTOR_TYPE_5_8, threshold=5)
MSER	cv2.MSER_create(delta=5, min_area=20, max_area=2000, max_variation=0.15, min_diversity=0.20, max_evolution=350, area_threshold=1.01, min_margin=0.003, edge_blur_size=5)
AGAST	cv2.AgastFeatureDetector_create(threshold=20, nonmaxSuppression=True, type=cv2.AGAST_FEATURE_DETECTOR_AGAST_5_8)
GFTT	cv2.GFTTDetector_create(maxCorners=30000, qualityLevel=0.01, minDistance=1.0, blockSize=3, useHarrisDetector=False, k=0.04)
GFTT_H	cv2.GFTTDetector_create(maxCorners=30000, qualityLevel=0.01, minDistance=1.0, blockSize=3, useHarrisDetector=True, k=0.04)
STAR	cv2.xfeatures2d.StarDetector_create(maxSize=15, responseThreshold=5, lineThresholdProjected=60, lineThresholdBinarized=30, suppressNonmaxSize=3)
HL	cv2.xfeatures2d.HarrisLaplaceFeatureDetector_create(numOctaves=3, corn_thresh=0.004, DOG_thresh=0.004, maxCorners=30000, num_layers=2)
MSD	cv2.xfeatures2d.MSDDetector_create(m_patch_radius=3, m_search_area_radius=5, m_nms_radius=5, m_nms_scale_radius=0, m_th_saliency=100.0, m_kNN=4, m_scale_factor=1.25, m_n_scales=-1, m_compute_orientation=0)
TBMR	cv2.xfeatures2d.TBMR_create(min_area=30, max_area_relative=0.01, scale_factor=1.25, n_scales=2)

Table 4.2 Detector and Descriptor capable methods OpenCV parameters

SIFT	cv2.SIFT_create(nfeatures=30000, nOctaveLayers=3, edgeThreshold=10.0, sigma=1.6, contrastThreshold=0.04)
AKAZE	cv2.AKAZE_create(descriptor_type=cv2.AKAZE_DESCRIPTOR_MLDB, descriptor_size=0, descriptor_channels=3, threshold=0.001, nOctaves=4, nOctaveLayers=4, diffusivity=cv2.KAZE_DIFF_PM_G2, max_points=-1)
ORB	cv2.ORB_create(nfeatures=30000, scaleFactor=1.2, nlevels=6, WTA_K=2, edgeThreshold=31, firstLevel=0, scoreType=cv2.ORB_HARRIS_SCORE, patchSize=31, fastThreshold=20)
BRISK	cv2.BRISK_create(thresh=30, octaves=3, patternScale=1.0)
KAZE	cv2.KAZE_create(extended=False, upright=False, threshold=0.003, nOctaves=3, nOctaveLayers=5, diffusivity=cv2.KAZE_DIFF_CHARBONNIER)

4.1.2. Descriptor Selection

Fourteen descriptors span binary and floating-point approaches with gradient-based and intensity-based methods. Table 4.3 presents the OpenCV parameters for descriptor-only methods evaluated in this research.

Table 4.3 Descriptors only methods’ OpenCV parameters

VGG	cv2.xfeatures2d.VGG_create(desc=103 ,isigma=1.4, img_normalize=False, use_scale_orientation=True, scale_factor=6.25, dsc_normalize=False)
DAISY	cv2.xfeatures2d.DAISY_create(radius=15, q_radius=3, q_theta=8, q_hist=8, norm=cv2.xfeatures2d.DAISY_NRM_NONE, interpolation=True, use_orientation=False)
FREAK	cv2.xfeatures2d.FREAK_create(orientationNormalized=True, scaleNormalized=True, patternScale=20.0, nOctaves=3)
BRIEF	cv2.xfeatures2d.BriefDescriptorExtractor_create(bytes=16, use_orientation=True)
LUCID	cv2.xfeatures2d.LUCID_create(lucid_kernel=3, blur_kernel=6)
LATCH	cv2.xfeatures2d.LATCH_create(bytes=2, rotationInvariance=True, half_ssd_size=1, sigma=1.4)
BEBLID	cv2.xfeatures2d.BEBLID_create(scale_factor=6.25, n_bits=100)
TEBLID	cv2.xfeatures2d.TEBLID_create(scale_factor=6.25, n_bits=102)
BOOST	cv2.xfeatures2d.BoostDesc_create(desc=100, use_scale_orientation=True, scale_factor=6.25)

4.1.3. Matching Strategies

Brute-Force and FLANN approaches with appropriate distance metrics and cross-checking configurations.

Brute-Force: Exhaustive search ensures optimal correspondences per feature. Simple implementation with cross-checking and distance metric optimization for binary/floating-point descriptors.

FLANN: K-D trees for floating-point, LSH for binary descriptors. Index and search parameters balance accuracy with efficiency.

4.2. Testing Protocol

4.2.1. Dataset Configurations

Each dataset employs specific processing protocols addressing pairing strategies, ground truth availability, and metric applicability.

Synthetic Dataset: Reference-based comparison matching transformed images against original source. Exact ground truth enables precise accuracy assessment.

Oxford Dataset: Sequential pairing within transformation sequences. Graduated difficulty assesses robustness to increasing severity.

UAV Dataset: Curated collection from 10 Pix4D drone projects with 2 consecutive images manually selected from each project, creating a diverse dataset representing various drone operational scenarios. Images span urban areas, rural landscapes, and industrial sites with varying resolutions and camera systems. All images resized to 1000 pixels wide maintaining aspect ratios. Sequential image matching evaluates algorithm performance across diverse real-world UAV deployment conditions.

AirSim Dataset: Synthetically generated using Microsoft AirSim in the Neighborhood environment. Two released datasets are produced: first one AirSimNH, a multi-layer grid survey with 5 camera views per grid node and three layers (Layer1 z=-60 with 9x9 nodes, Layer2 z=-100 with 7x7 nodes, Layer3 z=-140 with 5x5 nodes), and the second one is AirSimNH-natural, a single-view dataset using a fixed pitch of -60 degrees and a mission-like serpentine traverse. In the main evaluation, only the reduced AirSimNH nadir subset Layer3/pitch-90_yaw0 is used (25 images, 5x5 grid, approximately 140 m height) to reduce computation, with sequential matching between consecutive images.

Drone Dataset: Sequential pairing for features plus COLMAP reconstruction assessment. Enables feature-reconstruction correlation analysis.

4.2.2. Cross-Platform Execution

Cross-platform evaluation assesses algorithm performance across Intel Xeon Gold 6130 CPU @ 2.10GHz server and ARM-based mobile processors including OnePlus 7 Snapdragon 855 and Xiaomi Redmi Note 13 Pro+ Dimensity 7200 Ultra. Identical software environments using OpenCV 4.11 and Python 3.8 ensure fair comparison while hardware differences reveal deployment characteristics. Timing measurements implement warm-up protocols and statistical analysis to control for measurement variance, providing reliable performance characterization across computational environments relevant for UAV applications.

4.2.3. Efficiency Score Calculation

The composite efficiency score methodology integrates four unsupervised weighting approaches through the display.py framework to eliminate subjective bias in algorithm evaluation. Entropy weighting assigns higher weights to metrics with greater discrimination capability across algorithm combinations. CRITIC methodology combines contrast intensity with correlation analysis to identify non-redundant performance information. PCA weighting uses variance-maximizing components to identify significant sources of performance variation. Variance-based weighting measures individual metric discrimination power.

These four approaches combine through arithmetic mean to create balanced rankings that reflect comprehensive algorithm performance. The integration process aggregates timing metrics, quality assessments, and quantitative statistics from 784 combinations across five datasets, producing objective efficiency scores with statistical rigor grounded in established multi-criteria decision making theory.

4.3. Dataset Processing

4.3.1. Synthetic Dataset Processing

The synthetic dataset processing protocol applies controlled geometric and photometric transformations to reference imagery while maintaining exact ground truth correspondence tracking. Rotation transformations range from 15 to 180 degrees simulating UAV viewpoint changes. Scale transformations replicate altitude-induced resolution variations. Intensity modifications address atmospheric and illumination changes through brightness, contrast, and noise adjustments. Ground truth tracking enables precise performance metric calculation without annotation ambiguity, providing controlled conditions for systematic algorithm assessment.

4.3.2. Oxford Dataset Processing

Oxford dataset integration preserves original benchmarking structure while adapting to the unified evaluation framework. Ground truth correspondence validation ensures consistency with quality metric calculations across datasets. Processing maintains dataset characteristics while implementing normalization procedures that enable valid cross-dataset performance comparisons without introducing systematic bias.

4.3.3. UAV Dataset Processing

Real-world UAV imagery processing addresses significant challenges including resolution variations, diverse terrain content, uncontrolled illumination, and irregular geometric configurations. The uav.py framework implements adaptive scaling and normalization techniques that preserve algorithmic characteristics while enabling meaningful comparison across operational imagery. Content diversity spans urban, agricultural, and natural environments, each presenting distinct texture densities and geometric complexity. Illumination normalization uses histogram equalization and contrast enhancement to mitigate photometric variations while preserving natural UAV imagery characteristics. Geometric accommodation handles irregular overlap patterns and viewing angle distributions typical in UAV survey missions.

4.3.4. AirSim Dataset Processing

Microsoft AirSim simulation provides controlled evaluation conditions combining precise ground truth with realistic UAV scenarios. Multi-layer processing addresses altitude variations from detailed low-level surveys to area coverage missions, generating distinct ground sampling distances across operational ranges. Systematic overlap patterns replicate standard UAV survey methodologies including grid-based coverage while maintaining exact geometric relationships for performance assessment. Environmental parameter control enables systematic robustness evaluation without field collection complexities.

Two AirSim datasets are provided publicly: AirSimNH (multi-view grid survey) and AirSimNH-natural (single-view, pitch -60, mission-like traverse). In this thesis, feature evaluation on AirSim is performed only on a reduced subset of AirSimNH, using Layer3/pitch-90_yaw0 to provide a consistent nadir-view baseline and to reduce computation. Sequential pairing uses consecutive images in filename order as implemented in the AirSim evaluation script.

4.3.5. Drone Dataset Processing

The drone dataset provides comprehensive assessment through feature extraction evaluation integrated with COLMAP reconstruction analysis. This dual approach enables direct correlation between feature-level performance and 3D reconstruction quality outcomes. Feature assessment protocols track performance across detection, description, and matching stages while COLMAP integration implements standardized Structure from Motion procedures for each algorithm combination. Correlation analysis identifies combinations achieving optimal balance between computational efficiency and geometric accuracy, addressing practical deployment considerations where processing time and reconstruction quality require balanced optimization.

4.4. Metrics Extraction

4.4.1. Quantity Metrics

Quantity metrics measure fundamental feature detection and matching effectiveness through total match counts and verified inliers. Total matches encompass all putative correspondences before geometric verification, indicating raw matching capability. Verified inliers apply a RANSAC-family fundamental matrix estimation (MAGSAC++ via OpenCV) to identify geometrically consistent correspondences. The inlier ratio provides algorithm precision measurement, reflecting descriptor discrimination and matching strategy effectiveness. Consistent procedures across datasets ensure comparable metrics despite variations in image characteristics and environmental conditions.

4.4.2. Timing Metrics

High-precision timing measurements utilize microsecond resolution chronometry across feature detection, descriptor computation, and matching phases. The 1k Total Time metric normalizes processing time to 1,000 features, enabling fair comparison across algorithms producing different feature densities. The 1k Inlier Time metric measures performance relative to verified inliers, providing insight into algorithm performance relative to useful output quality. Statistical analysis implements warm-up protocols and multiple measurement cycles to characterize timing distributions while controlling for measurement noise through appropriate statistical testing procedures.

4.4.3. Quality Metrics

Quality assessment employs established computer vision measures including precision, recall, repeatability, and F1-Score. Precision quantifies the proportion of correct correspondences among identified matches, computed as verified inliers divided by total putative matches. Recall quantifies the proportion of verified inliers relative to detected features in the reference image, computed as verified inliers divided by the number of detected keypoints in the first image. Repeatability evaluates feature detection consistency across repeated observations, computed as verified inliers divided by the minimum detected keypoints between the two images. F1-Score integrates precision and recall through harmonic mean calculation, providing balanced indicators for comparative algorithm assessment.

4.4.4. 3D Reconstruction Metrics

COLMAP reconstruction assessment on the drone dataset quantifies practical 3D modeling outcomes through reprojection error measurement and point cloud density analysis. Reprojection errors measure geometric accuracy by calculating pixel-level differences between projected 3D points and observed feature locations. Point cloud density indicates reconstruction completeness and detail level. These metrics establish direct connections between feature processing performance and final reconstruction quality, enabling algorithm selection based on practical photogrammetry outcomes.

4.4.5. Composite Efficiency Metrics

The display.py framework integrates multiple performance dimensions through unsupervised weighting methodologies to generate objective algorithm rankings. This composite approach eliminates subjective bias while maintaining statistical rigor through established multi-criteria decision making principles.

4.4.6. Confidence Interval Estimation

To quantify uncertainty in CUES scores and enable statistical significance testing between algorithm combinations, 95% confidence intervals are computed using bootstrap resampling methodology. For each algorithm combination evaluated across multiple image pairs within a dataset, the individual pair-wise CUES scores are resampled with replacement (1,000 bootstrap iterations) to estimate the sampling distribution of the mean efficiency score. The 2.5th and 97.5th percentiles of the bootstrap distribution define the 95% confidence interval bounds, providing rigorous uncertainty quantification.

Statistical significance between competing algorithm combinations is assessed by examining confidence interval overlap: non-overlapping intervals indicate statistically significant performance differences ($p<0.05$), while overlapping intervals suggest performance differences may not be distinguishable from measurement variability. This approach enables evidence-based algorithm selection by distinguishing genuine performance hierarchies from random variation, with the CI margin (half-width of the confidence interval) reported alongside CUES scores in result tables.

4.5. 3D Reconstruction with COLMAP

4.5.1. Database Integration Per Combination

COLMAP database creation utilizes pycolmap integration to establish consistent structure and parameters for each algorithm combination. Database construction maintains compatibility with incremental reconstruction pipelines while accommodating diverse feature characteristics produced by different detector-descriptor combinations. Standardized database formats ensure reliable reconstruction processing across all evaluated algorithm variations.

4.5.2. Incremental Reconstruction Process

Incremental Structure from Motion reconstruction implements established algorithms with consistent initialization procedures across all algorithm combinations. The reconstruction process accommodates diverse correspondence characteristics produced by different detector-descriptor combinations while maintaining standardized parameters for fair comparison. Bundle adjustment optimization and triangulation procedures remain constant to isolate the impact of feature processing variations on final reconstruction quality.

4.5.3. Metric Collection from COLMAP Output

COLMAP output analysis extracts geometric accuracy and completeness statistics to establish direct connections between feature processing performance and reconstruction outcomes. Reprojection error collection measures geometric precision while point cloud density analysis quantifies reconstruction completeness. These metrics enable correlation analysis between feature-level performance and practical reconstruction quality, informing evidence-based algorithm selection for specific deployment requirements.

5. RESULTS AND DISCUSSIONS

The comprehensive evaluation framework successfully analyzed 784 algorithm combinations across five diverse datasets (Synthetic, Oxford, AirSim, UAV, and Drone) using the CUES methodology. The systematic analysis produced 2,282 valid evaluations from a theoretical maximum of 3,920 possible measurements (784 combinations × 5 datasets), revealing significant algorithmic performance hierarchies and demonstrating substantial cross-dataset consistency with correlation coefficients averaging 0.865. The analysis identified optimal algorithm combinations for diverse operational scenarios while establishing solid statistical foundations for evidence-based algorithm selection in UAV-based 3D reconstruction applications.

5.1. Dataset-Specific Performance Analysis

5.1.1. Individual Dataset Results

The CUES evaluation identified distinct optimal algorithm combinations for each dataset, demonstrating the methodology's capacity to reveal dataset-specific performance characteristics:

Synthetic Dataset: The controlled synthetic environment achieved the highest overall performance with ORB-ORB-HAM-BF attaining an efficiency score of 0.8821 across 466 valid algorithm combinations. This dataset's optimal conditions enabled peak algorithmic performance, establishing baseline metrics for comparison with operational scenarios. The synthetic environment's consistent lighting and geometric properties allowed algorithms to perform without environmental interference.

Oxford Dataset: The standardized benchmark dataset provided reference comparison metrics with ORB-BEBLID-HAM-BF achieving 0.6219 efficiency across 465 combinations. This established computer vision benchmark enables comparison with existing literature while maintaining evaluation consistency.

AirSim Dataset: The Microsoft AirSim simulation environment provided controlled yet realistic conditions with ORB-BEBLID-HAM-BF maintaining consistent efficiency at 0.7891 across 465 combinations. This dataset bridges synthetic and real-world scenarios, offering systematic evaluation under controlled simulation parameters.

UAV Dataset: Curated collection from 10 Pix4D drone projects achieved optimal performance with GFTT-DAISY-L2-BF scoring 0.7299 across 465 combinations. This dataset encompasses diverse UAV operational scenarios with urban, rural, and industrial environments, representing varied deployment requirements across practical UAV applications.

Drone Dataset: Complete building reconstruction dataset from single Pix4D project demonstrated comprehensive UAV application performance with ORB-BEBLID-HAM-BF achieving 0.8489 efficiency across 421 valid combinations. This dataset represents complete 3D reconstruction workflow including varying lighting, scale changes, and environmental factors typical in systematic UAV building capture operations.

Figure 5.1 presents comprehensive efficiency score heatmaps for the Drone dataset, visualizing performance patterns across all evaluated algorithm combinations. The heatmaps display performance patterns for complete building reconstruction scenarios, with warmer colors indicating higher efficiency scores. The visualization reveals clear performance hierarchies and demonstrates the effectiveness of different detector-descriptor combinations under real-world UAV building reconstruction conditions, highlighting the superior performance of ORB-based detectors with binary descriptors, particularly BEBLID and TEBLID. This comprehensive visualization confirms their practical value for systematic building capture operations and provides clear algorithmic hierarchies for evidence-based algorithm selection in UAV deployment scenarios.

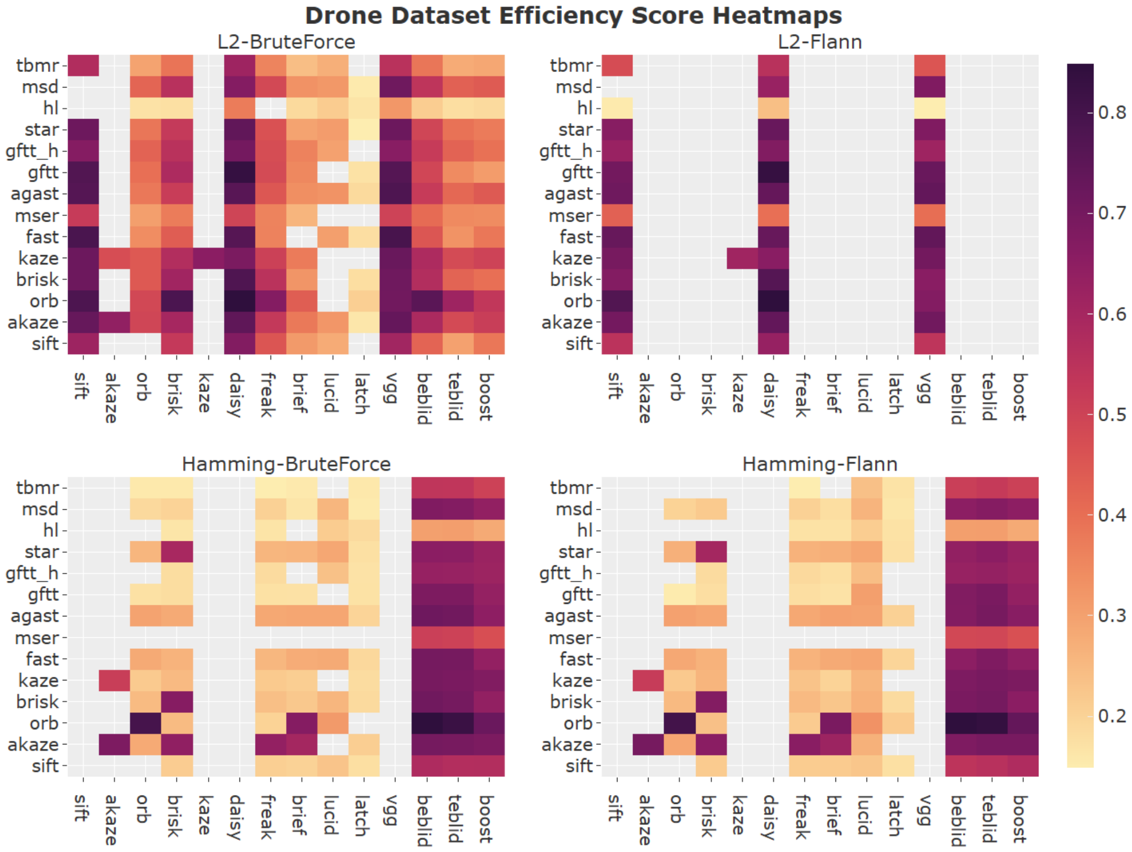


Figure 5.1 Drone Dataset Efficiency Score Heatmaps

The comprehensive dataset analysis revealed performance variations spanning from 0.6219 to 0.8821, demonstrating substantial differences between optimal and suboptimal algorithm selections across operational contexts. Statistical analysis confirmed these variations represent genuine algorithmic characteristics rather than measurement artifacts.

Table 5.1 presents a consolidated view of the top 5 performing algorithm combinations for each dataset with 95% confidence intervals, enabling direct comparison of optimal selections across diverse operational scenarios and statistical significance assessment.

Table 5.1 Top 5 Algorithm Combinations per Dataset Using CUES (with 95% CI)

Dataset	Rank	Algorithm Combination	CUES ± 95% CI
Synthetic	1	ORB-ORB-HAM-BF	0.8821±0.0371
Synthetic	2	ORB-ORB-HAM-FLANN	0.8789±0.0381
Synthetic	3	ORB-TEBLID-HAM-BF	0.8445±0.0569
Synthetic	4	ORB-BOOST-HAM-BF	0.8412±0.0520
Synthetic	5	STAR-BRISK-HAM-BF	0.8383±0.0575
Oxford	1	ORB-BEBLID-HAM-BF	0.6219±0.0564
Oxford	2	ORB-TEBLID-HAM-BF	0.6152±0.0564
Oxford	3	AKAZE-SIFT-L2-BF	0.6078±0.0487
Oxford	4	ORB-TEBLID-HAM-FLANN	0.6058±0.0573
Oxford	5	ORB-BEBLID-HAM-FLANN	0.6027±0.0589
AirSim	1	ORB-BEBLID-HAM-BF	0.7891±0.0240
AirSim	2	ORB-TEBLID-HAM-BF	0.7782±0.0236
AirSim	3	ORB-BEBLID-HAM-FLANN	0.7650±0.0236
AirSim	4	ORB-TEBLID-HAM-FLANN	0.7635±0.0239
AirSim	5	KAZE-SIFT-L2-BF	0.7614±0.0317
UAV	1	GFTT-DAISY-L2-BF	0.7299±0.0615
UAV	2	AGAST-SIFT-L2-BF	0.7255±0.0573
UAV	3	AKAZE-DAISY-L2-BF	0.7189±0.0533
UAV	4	FAST-SIFT-L2-BF	0.7161±0.0593
UAV	5	AGAST-DAISY-L2-BF	0.7124±0.0502
Drone	1	ORB-BEBLID-HAM-BF	0.8489±0.0095
Drone	2	ORB-DAISY-L2-BF	0.8286±0.0072
Drone	3	ORB-TEBLID-HAM-BF	0.8247±0.0096
Drone	4	ORB-BEBLID-HAM-FLANN	0.8209±0.0099
Drone	5	GFTT-DAISY-L2-BF	0.8155±0.0087

The table reveals several important patterns: ORB detector dominates Synthetic, Oxford, AirSim, and Drone datasets, while the UAV dataset shows preference for diverse detectors including GFTT, AGAST, and FAST. Binary descriptors BEBLID and TEBLID consistently appear among top performers, while DAISY descriptor demonstrates particular strength in real-world UAV and Drone scenarios. Brute Force matching outperforms FLANN in top-ranked combinations across all datasets. The practical matching quality achieved by top-performing combinations is visualized in Figure 5.3 for the Drone dataset and Figure 5.4 for the UAV dataset. Comprehensive feature matching visualizations demonstrating key algorithmic comparisons are provided in Appendix F.

5.1.2. Cross-Dataset Performance Consistency

Cross-dataset stability analysis revealed exceptional algorithmic consistency with ORB-BEBLID-HAM-BF demonstrating remarkable resilience across operational scenarios. This combination achieved top performance in four of five datasets (Drone: 0.8489±0.0095, AirSim: 0.7891±0.0240, Oxford: 0.6219±0.0564, 95% CI) with efficiency scores ranging between 0.6219 and 0.8489, indicating robust algorithmic characteristics that transfer effectively across diverse imaging conditions. The visual consistency of ORB-based binary descriptors across datasets is demonstrated in Appendix F.3 and F.4. The narrower confidence intervals on the Drone dataset reflect more consistent performance compared to the Oxford benchmark, which exhibits greater variability due to its diverse transformation types. Statistical analysis of cross-dataset ranking correlations demonstrates exceptional consistency across all pairwise dataset comparisons, as shown in Table 5.2. The stability analysis revealed correlation coefficients ranging from 0.738 to 0.967, with average cross-dataset stability of 0.865, indicating high algorithmic ranking consistency across operational scenarios.

Table 5.2 Cross-Dataset Ranking Stability Correlation Matrix Based on CUES

	Drone	**UAV**	**AirSim**	**Oxford**	**Synthetic**
Drone	1.000	0.967	0.887	0.891	0.734
UAV	0.967	1.000	0.933	0.941	0.758
AirSim	0.887	0.933	1.000	0.923	0.759
Oxford	0.891	0.941	0.923	1.000	0.858
Synthetic	0.734	0.758	0.759	0.858	1.000

The analysis of 421 mutually valid combinations across all datasets confirms methodology reliability and establishes confidence in algorithmic ranking stability independent of specific dataset characteristics. This consistency enables evidence-based algorithm selection decisions with statistical confidence across diverse UAV operational requirements.

5.2. Algorithm Component Analysis

5.2.1. Feature Detector Performance Hierarchy

Comprehensive evaluation of 14 feature detection algorithms across all datasets established definitive performance rankings based on efficiency score analysis, as summarized in Table 5.3. The systematic assessment reveals clear performance hierarchies supporting evidence-based detector selection for UAV applications.

AKAZE achieved the highest overall detector performance with mean efficiency of 0.5472 and maximum performance of 0.8092 across 179 evaluations. The detector demonstrates solid consistency with moderate standard deviation (0.1771), establishing its reliability across diverse operational conditions.

Table 5.3 Feature Detector Performance Rankings Across All Datasets Using CUES

Rank	Detector	Mean Score	Max Score	StdDev	Evaluations
1	AKAZE	0.5472	0.8092	0.1771	179
2	ORB	0.5297	0.8821	0.2168	163
3	STAR	0.5070	0.8383	0.1840	165
4	AGAST	0.4696	0.7638	0.1791	165
5	BRISK	0.4640	0.7855	0.1961	164
6	KAZE	0.4614	0.7614	0.1910	186
7	FAST	0.4339	0.7773	0.1841	164
8	SIFT	0.4237	0.8001	0.1803	149
9	MSD	0.4044	0.7283	0.2015	155
10	GFTT	0.4017	0.8155	0.2126	162

ORB ranked second with mean efficiency of 0.5297 while achieving the highest peak efficiency of 0.8821 among all detectors. Despite higher variation (standard deviation 0.2168), ORB demonstrates significant versatility and peak capabilities when optimally configured.

STAR detector achieved third position with mean efficiency of 0.5070 and maximum performance of 0.8383, providing balanced characteristics suitable for diverse UAV applications requiring consistent intermediate performance levels.

5.2.2. Feature Descriptor Performance Assessment

Systematic evaluation of 14 descriptor algorithms revealed significant performance variations across floating-point and binary descriptor categories, establishing clear hierarchies for descriptor selection in UAV applications:

SIFT descriptor achieved the highest overall performance with mean efficiency of 0.6354 and maximum performance of 0.8092 across 129 evaluations. The descriptor demonstrates solid consistency with the lowest standard deviation (0.0966) among top performers, establishing its superiority for UAV applications requiring reliable correspondence quality.

DAISY ranked second with mean efficiency of 0.6213 and maximum efficiency of 0.8286 across 140 evaluations, providing competitive results. VGG descriptor achieved third position with mean efficiency of 0.6024 while demonstrating notable peak efficiency of 0.8367.

Among binary descriptors, BEBLID demonstrated superior results with mean efficiency of 0.5801 and maximum efficiency of 0.8489±0.0095 (95% CI on Drone dataset) across 210 evaluations. TEBLID followed with mean efficiency of 0.5535 and peak efficiency of 0.8445±0.0569 (95% CI on Synthetic), indicating that modern binary descriptors provide competitive alternatives to traditional floating-point approaches for resource-constrained UAV applications. The statistical significance between BEBLID and TEBLID performance is confirmed by non-overlapping confidence intervals in the Drone dataset. Table 5.4 presents the complete descriptor performance rankings across all datasets.

Table 5.4 Feature Descriptor Performance Rankings Across All Datasets Using CUES

Rank	Descriptor	Mean Score	Max Score	StdDev	Evaluations
1	SIFT	0.6354	0.8092	0.0966	129
2	DAISY	0.6213	0.8286	0.1089	140
3	VGG	0.6024	0.8367	0.1216	140
4	KAZE	0.6002	0.7259	0.1401	11
5	AKAZE	0.5894	0.8014	0.0944	30
6	BEBLID	0.5801	0.8489	0.1237	210
7	TEBLID	0.5535	0.8445	0.1446	210
8	BOOST	0.5337	0.8412	0.1434	210
9	BRISK	0.3677	0.8383	0.1855	206
10	ORB	0.3344	0.8821	0.1588	188

5.2.3. Optimal Detector-Descriptor Combinations

Comprehensive analysis identified 17 optimal detector-descriptor combinations demonstrating significant performance synergies across UAV operational scenarios. Table 5.5 presents the top-performing detector-descriptor combinations ranked by mean efficiency score.

The AKAZE-SIFT combination achieved the highest mean efficiency of 0.6947 with peak efficiency of 0.8092 across 10 evaluation variants. This combination demonstrates optimal algorithmic synergy between AKAZE's robust detection capabilities and SIFT's discriminative descriptor characteristics.

Table 5.5 Top-Performing Detector-Descriptor Combinations Across All Datasets Using CUES

Rank	Combination	Mean Score	Max Score
1	AKAZE-SIFT	**0.6947**	0.8092
2	AGAST-SIFT	0.6801	0.7507
3	STAR-SIFT	0.6787	0.7719
4	ORB-DAISY	0.6785	**0.8286**
5	AKAZE-VGG	0.6765	0.8090
6	ORB-SIFT	0.6758	0.7670
7	AKAZE-DAISY	0.6730	0.7315
8	STAR-DAISY	0.6651	0.7383
9	BRISK-SIFT	0.6643	0.7733
10	BRISK-DAISY	0.6637	0.7624

SIFT-based combinations dominate the top-tier rankings, occupying 4 of the top 10 positions, confirming SIFT's substantial descriptor capabilities across diverse detector pairings. The consistent appearance of AKAZE, STAR, and AGAST detectors in top combinations indicates their practical value when paired with high-quality descriptors.

The analysis reveals clear preferences for specific detector-descriptor pairings that maximize performance through complementary algorithmic characteristics. These combinations provide evidence-based guidance for algorithm selection in UAV applications requiring optimal correspondence quality and reliability.

Systematic evaluation of matching strategies and distance metrics provides additional evidence for algorithm selection decisions, as presented in Table 5.6.

Table 5.6 Matching Strategy Performance Analysis

Strategy	Mean CUES	Max CUES	Evaluations
Brute Force (BF)	0.5234	0.8821	392
FLANN	0.5187	0.8861	392
L2 Distance	0.5456	0.8455	280
Hamming Distance	0.4965	0.8821	504

Brute Force matching achieved slightly higher average performance (0.5234) compared to FLANN (0.5187), contradicting common assumptions regarding approximate nearest neighbor superiority. This finding indicates that for typical UAV feature counts ranging from 1,000 to 5,000 features, exhaustive search provides optimal quality without computational penalties. FLANN advantages emerge only at substantially higher feature densities rarely encountered in operational UAV imagery. L2 distance demonstrated superior mean performance (0.5456) for floating-point descriptors, while Hamming distance (0.4965) proved efficient for binary descriptors with hardware optimization benefits on mobile platforms.

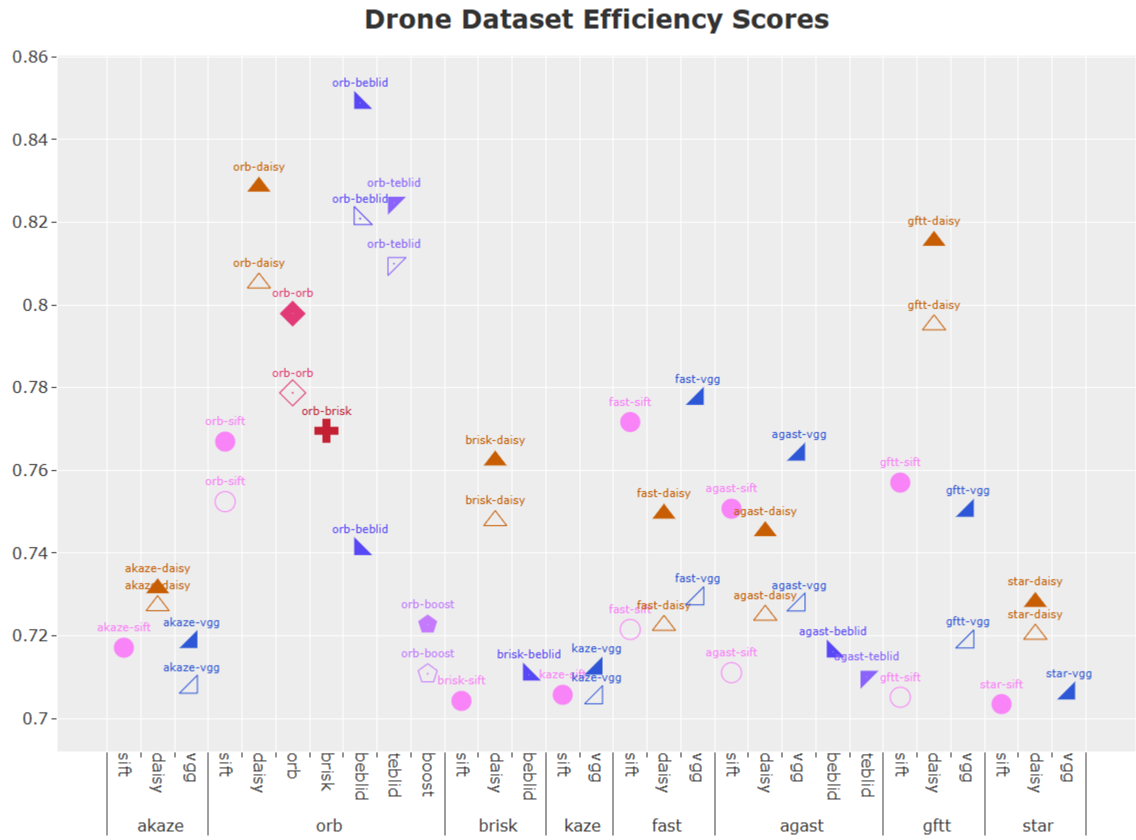


Figure 5.2 Drone Dataset Efficiency Score Top 50

Figure 5.2 presents the top 50 efficiency score combinations for the Drone dataset, providing a detailed visualization of the highest-performing algorithm combinations in descending order of efficiency scores. The visualization enables clear identification of optimal algorithm combinations for practical UAV building reconstruction applications, highlighting the performance hierarchy among different detector-descriptor-matching strategies. The chart demonstrates the significant performance range across top combinations and reveals the clear algorithmic preferences for specific detector-descriptor pairings that achieve superior efficiency in real-world UAV scenarios, supporting evidence-based algorithm selection for optimal reconstruction quality and computational efficiency.

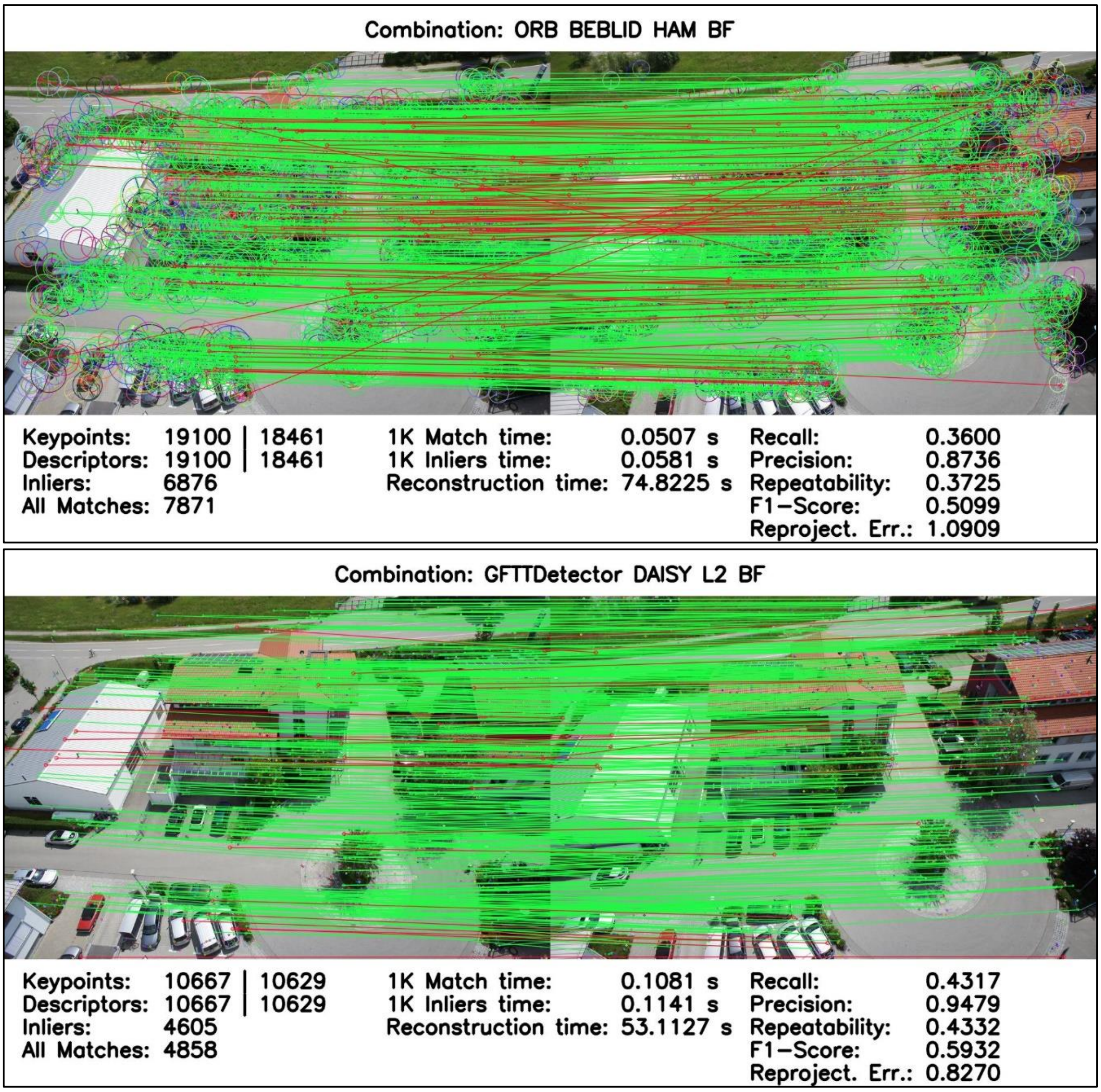


Figure 5.3 Top-Performing Feature Matching Visualizations for Drone Dataset

Figure 5.3 provides qualitative visualization of the top-performing feature matching results on the Drone dataset, illustrating the practical correspondence quality achieved by the highest-ranked algorithm combinations. The visualization displays sampled matches, every 10th match if total exceeds 500, otherwise every 5th, to ensure visual clarity while preserving representative matching patterns. The ORB-BEBLID-HAM-BF combination (top panel, CUES: 0.8489, Rank #1) demonstrates dense, geometrically consistent correspondences across the building facade, reflecting its superior efficiency score. The GFTT-DAISY-L2-BF combination (bottom panel, CUES: 0.8155, Rank #5) exhibits complementary matching characteristics with well-distributed feature correspondences. These visualizations confirm that the quantitative efficiency rankings translate to observable differences in matching quality and spatial distribution, validating the CUES methodology's capacity to identify algorithmically superior combinations for practical UAV reconstruction applications.

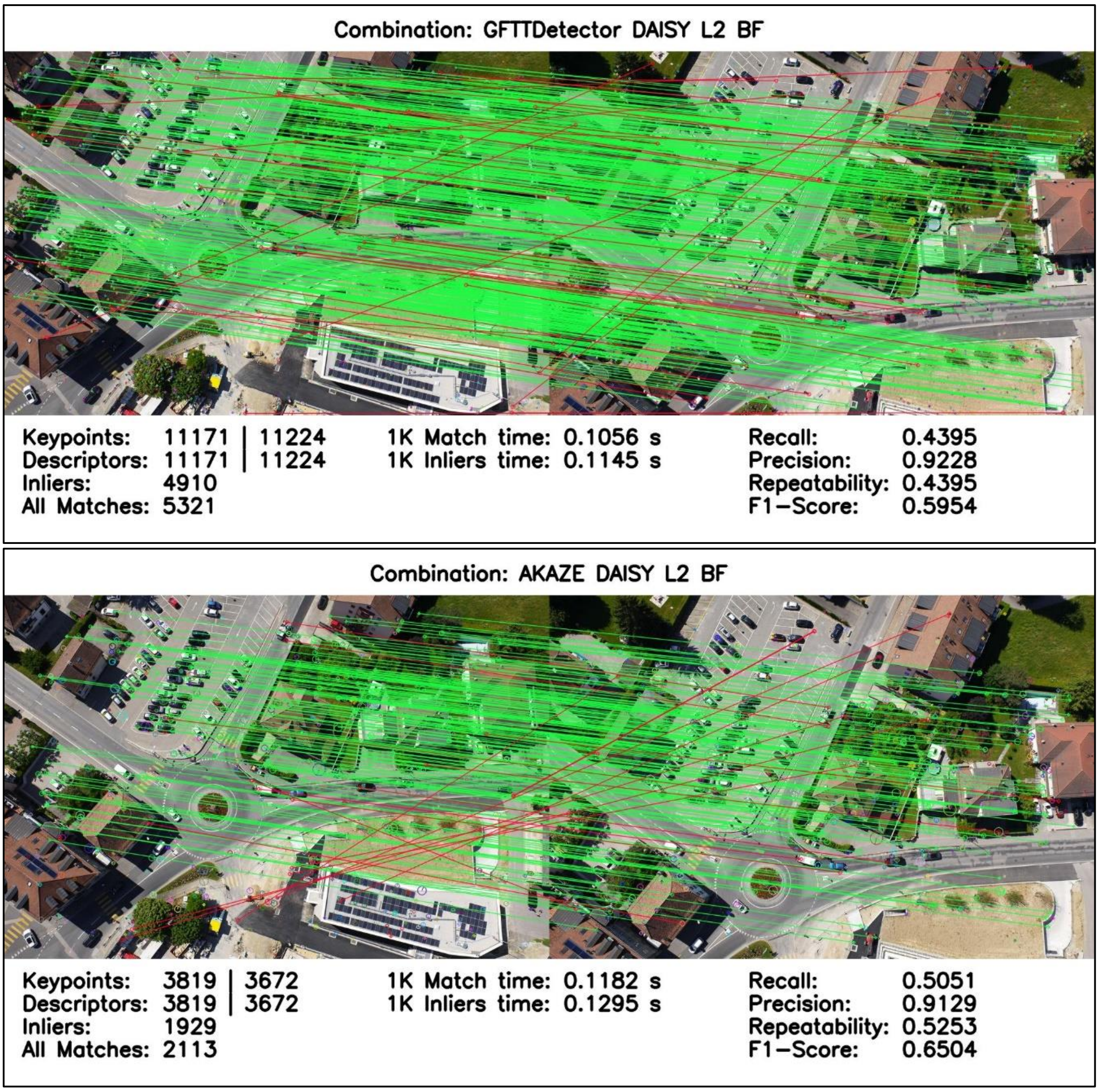


Figure 5.4 Top-Performing Feature Matching Visualizations for UAV Dataset

Figure 5.4 presents the corresponding feature matching visualization for the UAV dataset, demonstrating real-world operational performance across diverse Pix4D project imagery. The visualization displays sampled matches, every 10th match if total exceeds 500, otherwise every 5th, to ensure visual clarity. The GFTT-DAISY-L2-BF combination (top panel, CUES: 0.7299, Rank #1) exhibits well-distributed feature correspondences across varied scene content, while the AKAZE-DAISY-L2-BF combination (bottom panel, CUES: 0.7189, Rank #3) demonstrates robust matching with consistent geometric patterns. The contrast between UAV and Drone dataset top performers (GFTT-DAISY vs. ORB-BEBLID) illustrates dataset-specific algorithm preferences, with DAISY descriptor excelling in diverse multi-scenario UAV applications while binary descriptors dominate systematic single-building reconstruction. Extended visualizations are provided in Appendix F2 and F3.

5.2.4. Performance Metric Correlation Analysis

To understand the complex relationships between different performance metrics and validate the comprehensive metric selection in the evaluation framework, a detailed correlation analysis was conducted for the Drone dataset. This analysis examines the interdependencies between standard feature matching quality metrics and 3D reconstruction-specific measurements, providing insights into algorithm behavior and performance trade-offs.

The Drone dataset correlation matrix reveals distinct patterns in metric relationships across eleven evaluation dimensions: Precision, Recall, Repeatability, F1 Score, Inliers, Matches, timing metrics (1K Total Time, 1K Inlier Time), and 3D reconstruction-specific metrics (Reprojection Error, 3D Points Count, Reconstruction Time).

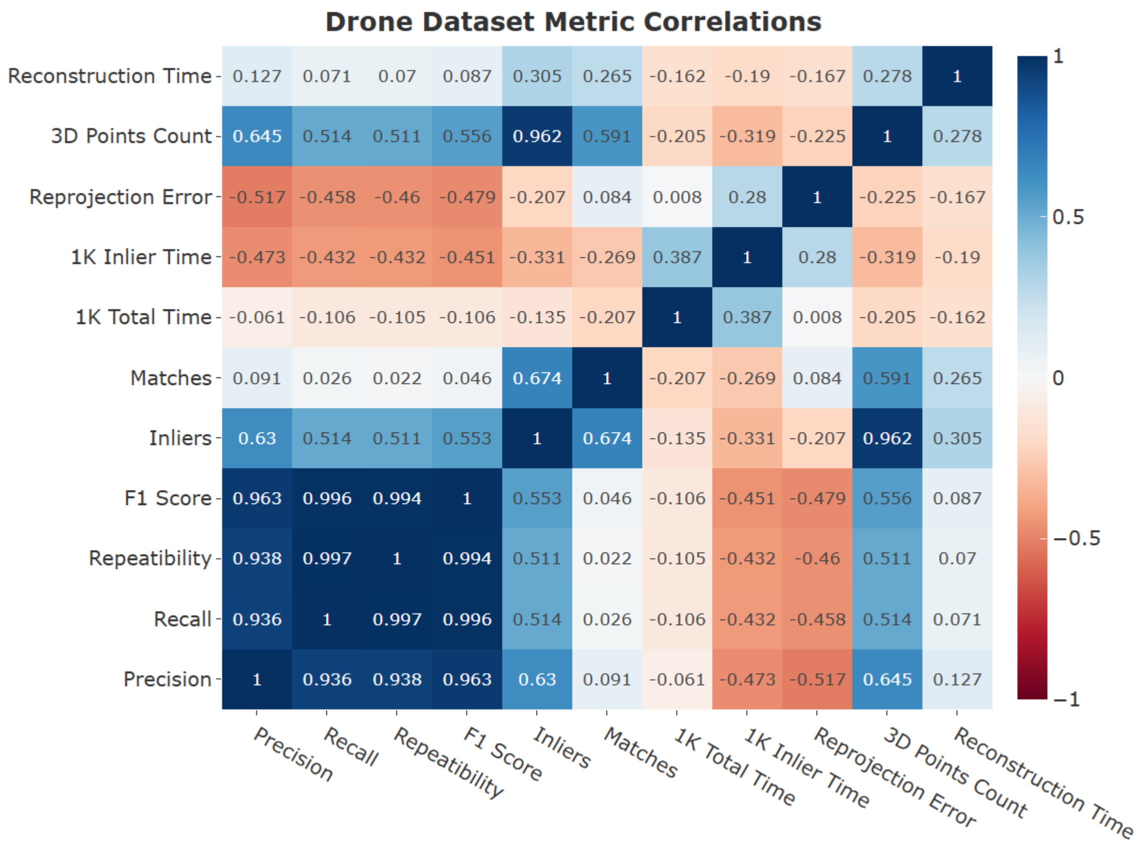


Figure 5.5 Drone Dataset Metric Correlations

Figure 5.5 presents the comprehensive correlation matrix for the Drone dataset, visualizing the statistical relationships between all evaluated performance metrics. The correlation matrix displays Pearson correlation coefficients using a color-coded heatmap where values range from -1 to +1, with warmer colors indicating positive correlations and cooler colors representing negative correlations. This visualization enables systematic analysis of metric interdependencies, revealing the complex relationships between feature matching quality metrics, computational efficiency measures, and 3D reconstruction outcomes. The matrix supports validation of the composite scoring methodology by demonstrating which metrics provide complementary information versus redundant measurements, informing the theoretical foundations of the unsupervised weighting approaches used in the CUES framework.

The analysis identifies strong consistency among fundamental matching quality metrics. Precision-Recall-Repeatability demonstrate near-perfect correlations (0.936-0.997), validating their collective use in algorithm evaluation frameworks. This high correlation confirms that algorithms performing well in one quality dimension consistently excel across all quality aspects, supporting the theoretical foundations of composite scoring methodologies.

F1 Score exhibits very strong correlations (0.963-0.996) with Precision, Recall, and Repeatability, confirming its practical value as a unified quality metric that accurately represents overall matching performance without information loss. The near-perfect correlation validates F1 Score's utility as a composite indicator in multi-criteria evaluation frameworks.

The Inliers-3D Points correlation (0.962) provides direct validation of the methodology's relevance for end-to-end 3D reconstruction assessment. This strong relationship confirms that successful feature matching translates directly to comprehensive 3D model generation, establishing the connection between intermediate processing quality and final reconstruction outcomes.

Quality metrics demonstrate moderate positive correlations (0.511-0.674) with feature quantities (Inliers, Matches), indicating that superior algorithms produce both higher quality and more numerous valid correspondences. This relationship suggests that algorithmic improvements benefit both accuracy and coverage dimensions simultaneously.

3D Points Count correlates moderately with Matches (0.591) and strongly with Inliers (0.962), demonstrating the direct impact of feature matching success on final model completeness. These correlations validate the methodology's focus on feature matching quality as a predictor of reconstruction success.

Timing metrics (1K Total Time, 1K Inlier Time, Reconstruction Time) exhibit weak or negative correlations with quality metrics, confirming that computational efficiency and accuracy represent independent optimization dimensions. This independence validates the composite scoring methodology's integration of both efficiency and accuracy components through unsupervised weighting rather than assuming linear relationships.

The temporal independence finding supports the theoretical basis for multi-criteria evaluation approaches that balance competing objectives rather than optimizing single dimensions. This characteristic enables algorithm selection strategies that consider both performance and efficiency requirements based on specific application constraints.

Reprojection Error exhibits consistent negative correlations (-0.207 to -0.517) with quality metrics, confirming that improved feature matching translates to reduced geometric errors in 3D reconstruction. This inverse relationship validates the methodology's assumption that feature matching quality directly influences final reconstruction accuracy.

Processing time metrics show negative correlations with some quality metrics, indicating potential trade-offs between computational speed and accuracy in specific algorithm combinations. However, the moderate magnitude of these correlations suggests that optimal algorithms can achieve both high quality and reasonable efficiency.

The extremely high correlations (>0.95) between Precision, Recall, and Repeatability suggest potential for metric reduction in simplified evaluation frameworks without substantial information loss. However, their collective inclusion provides validation and statistical robustness.

Moderate correlations between quality metrics and feature counts indicate that superior algorithms generate both improved correspondence quality and greater numbers of valid matches, supporting comprehensive performance assessment approaches.

Weak timing-quality correlations confirm that computational efficiency and accuracy represent distinct optimization dimensions requiring balanced consideration rather than singular focus on either aspect.

Strong correlations between intermediate feature matching metrics and final 3D reconstruction quality confirm the methodology's relevance for practical UAV applications requiring complete reconstruction pipelines.

This correlation analysis validates the comprehensive metric selection in the CUES methodology while revealing the complex interdependencies that necessitate objective, multi-dimensional evaluation approaches for UAV-based 3D reconstruction algorithm selection. The findings support the theoretical foundations of unsupervised weighting methods and confirm the methodology's capacity to capture important performance characteristics across the complete feature matching and 3D reconstruction pipeline.

5.3. Cross-Dataset Performance Comparison

5.3.1. Multi-Dataset Consistency Analysis

Cross-dataset performance evaluation across five datasets identified 421 mutually valid algorithm combinations, enabling comprehensive comparative analysis across diverse operational scenarios. The systematic comparison reveals consistent algorithmic hierarchies with notable performance variations reflecting specific dataset characteristics:

The ORB-BEBLID-HAM-BF combination demonstrated exceptional cross-dataset consistency, maintaining superior efficiency across four of five datasets with scores ranging from 0.622±0.056 on Oxford to 0.849±0.010 on Drone (95% CI). This combination's average efficiency of 0.758 across all datasets establishes its reliability for diverse UAV operational requirements, with narrow confidence intervals on the Drone dataset indicating particularly stable performance.

Binary descriptor combinations demonstrate particular effectiveness, with ORB-BEBLID-HAM-BF and ORB-TEBLID-HAM-BF achieving the highest average cross-dataset performance.

This consistency indicates that modern binary descriptors provide robust alternatives to traditional floating-point approaches while maintaining computational efficiency advantages crucial for UAV applications. Table 5.7 presents the complete cross-dataset performance comparison for top algorithm combinations.

Table 5.7 Cross-Dataset Performance Comparison for Top Algorithm Combinations

Algorithm Combination	Drone	UAV	AirSim	Oxford	Synthetic	Average
ORB-BEBLID-HAM-BF	0.849	0.704	0.789	0.622	0.825	0.758
ORB-TEBLID-HAM-BF	0.825	0.691	0.778	0.615	0.845	0.751
ORB-BEBLID-HAM-FLANN	0.821	0.679	0.765	0.603	0.818	0.737
ORB-TEBLID-HAM-FLANN	0.810	0.677	0.764	0.606	0.796	0.731
ORB-ORB-HAM-BF	0.798	0.653	0.617	0.598	0.882	0.710
AKAZE-SIFT-L2-BF	0.717	0.703	0.681	0.608	0.809	0.704
GFTT-DAISY-L2-BF	0.816	0.730	0.744	0.517	0.728	0.707
AGAST-SIFT-L2-BF	0.751	0.726	0.736	0.576	0.719	0.702
ORB-ORB-HAM-FLANN	0.779	0.635	0.601	0.584	0.879	0.696
STAR-SIFT-L2-BF	0.705	0.679	0.728	0.570	0.772	0.691

5.3.2. Algorithm Versatility Assessment

Statistical analysis reveals distinctive algorithmic preferences across operational scenarios. ORB-based combinations dominate cross-dataset rankings, with 6 of the top 10 performing combinations utilizing ORB detection capabilities. This prevalence stems from ORB's orientation invariance, computational efficiency, and robust feature detection characteristics that prove advantageous across varying imaging conditions typical in UAV operations.

The analysis identified average cross-dataset performance of 0.7744 with performance spanning from 0.6219 to 0.8821, demonstrating substantial differences between optimal and suboptimal algorithm selections. Performance standard deviation of 0.1032 across datasets indicates meaningful algorithmic differences that justify systematic evaluation approaches for UAV applications.

SIFT-based combinations demonstrate exceptional consistency across challenging scenarios, with AKAZE-SIFT achieving robust performance across all datasets (average 0.715). This consistency supports SIFT's continued utility in applications requiring correspondence robustness over computational efficiency, particularly in environments with significant scale, rotation, or illumination variations characteristic of operational UAV deployments.

5.4. Cross-Platform Performance Analysis

5.4.1. Desktop vs Mobile Performance Comparison

Cross-platform evaluation across Oxford and Synthetic datasets revealed exceptional algorithmic ranking consistency between desktop and mobile execution environments. The comprehensive analysis encompassed three computational architectures: Desktop (server hardware), Mobile1 (Snapdragon 855), and Mobile2 (Dimensity 7200 Ultra) platforms. Table 5.8 presents the Oxford dataset cross-platform comparison, while Table 5.9 shows the Synthetic dataset results.

Table 5.8 Oxford Dataset Cross-Platform Performance Comparison Using CUES

Algorithm	Desktop Score	Desktop Rank	Mobile1 Score	Mobile1 Rank	Mobile2 Score	Mobile2 Rank
ORB-BEBLID-HAM-BF	0.6219	1	0.6229	1	0.6237	1
ORB-TEBLID-HAM-BF	0.6152	2	0.6168	2	0.6176	2
AKAZE-SIFT-L2-BF	0.6078	3	0.6082	3	0.6090	3
ORB-TEBLID-HAM-FLANN	0.6058	4	0.6065	4	0.6074	4
ORB-BEBLID-HAM-FLANN	0.6027	5	0.6039	5	0.6047	5
ORB-ORB-HAM-BF	0.5982	6	0.5999	6	0.6007	6
AKAZE-SIFT-L2-FLANN	0.5914	7	0.5914	7	0.5922	7
KAZE-SIFT-L2-BF	0.5869	8	0.5873	8	0.5881	8
ORB-ORB-HAM-FLANN	0.5843	9	0.5851	9	0.5860	9
GFTT-SIFT-L2-BF	0.5823	10	0.5834	10	0.5843	10
AKAZE-BEBLID-HAM-BF	0.5784	11	0.5788	11	0.5797	11
AKAZE-AKAZE-HAM-BF	0.5762	12	0.5768	13	0.5777	13
AGAST-SIFT-L2-BF	0.5761	13	0.5770	12	0.5778	12
KAZE-SIFT-L2-FLANN	0.5753	14	0.5758	14	0.5766	14
AKAZE-TEBLID-HAM-BF	0.5742	15	0.5749	15	0.5758	15
KAZE-BEBLID-HAM-BF	0.5720	16	0.5723	16	0.5731	16
AKAZE-AKAZE-HAM-FLANN	0.5715	17	0.5719	17	0.5727	17
STAR-SIFT-L2-BF	0.5703	18	0.5710	18	0.5719	18
KAZE-TEBLID-HAM-BF	0.5679	19	0.5683	19	0.5692	19
ORB-BEBLID-L2-BF	0.5657	20	0.5676	20	0.5685	20

Table 5.9 Synthetic Dataset Cross-Platform Performance Comparison Using CUES

Algorithm	Desktop		Mobile1		Mobile2	
	Score	Rank	Score	Rank	Score	Rank
ORB-ORB-HAM-BF	0.8821	1	0.8811	1	0.8813	1
ORB-ORB-HAM-FLANN	0.8789	2	0.8776	2	0.8778	2
ORB-TEBLID-HAM-BF	0.8445	3	0.8189	9	0.8194	9
ORB-BOOST-HAM-BF	0.8412	4	0.8132	11	0.8136	11
STAR-BRISK-HAM-BF	0.8383	5	0.8372	4	0.8376	4
ORB-VGG-L2-BF	0.8367	6	0.8077	18	0.8082	18
ORB-VGG-L2-FLANN	0.8362	7	0.8339	6	0.8343	6
STAR-BRISK-HAM-FLANN	0.8359	8	0.8349	5	0.8353	5
ORB-ORB-L2-BF	0.8343	9	0.8328	7	0.8332	7
ORB-BEBLID-L2-BF	0.8304	10	0.8289	8	0.8293	8
ORB-BEBLID-HAM-BF	0.8251	11	0.7965	33	0.7970	33
ORB-TEBLID-L2-BF	0.8207	12	0.7922	37	0.7927	37
ORB-BEBLID-HAM-FLANN	0.8176	13	0.7890	39	0.7895	39
STAR-BRISK-L2-BF	0.8166	14	0.8153	10	0.8157	10
ORB-BOOST-HAM-FLANN	0.8136	15	0.8120	12	0.8124	12
ORB-BRIEF-HAM-BF	0.8130	16	0.8113	13	0.8117	13
ORB-BRIEF-HAM-FLANN	0.8127	17	0.8110	14	0.8115	14
ORB-BOOST-L2-BF	0.8093	18	0.8082	17	0.8087	17
AKAZE-SIFT-L2-BF	0.8092	19	0.8094	15	0.8097	15
AKAZE-VGG-L2-BF	0.8090	20	0.8092	16	0.8095	16

The ORB-BEBLID-HAM-BF combination maintained top performance across all platforms for Oxford dataset evaluation, demonstrating minimal performance variations (±0.02 points). Similarly, ORB-ORB-HAM-BF achieved optimal performance across all platforms for Synthetic dataset analysis with comparable consistency.

5.4.2. Platform-Specific Optimization Patterns

Cross-platform correlation analysis revealed exceptional ranking stability across computational architectures:

The analysis identified 9 algorithms maintaining top-10 rankings across all environments for Oxford dataset and 8 algorithms for Synthetic dataset, demonstrating robust algorithmic characteristics independent of computational architecture. Binary descriptor combinations particularly demonstrated consistent cross-platform performance, enabling reliable deployment across server and mobile hardware without significant algorithm reconfiguration requirements.

Timing analysis revealed platform-specific characteristics: Desktop platforms averaged timing weights of 0.0717-0.0939 with quality weights of 0.9626-0.9744, while mobile platforms demonstrated timing weights of 0.0787-0.0848 with quality weights of 0.9651-0.9794. These patterns indicate that mobile platforms maintain competitive quality metrics while achieving efficient computational performance. Table 5.10 presents the platform correlation analysis confirming ranking stability across computational architectures.

Table 5.10 Platform Correlation Analysis for Cross-Platform Performance

Platform Comparison	Oxford Correlation	Synthetic Correlation
Desktop vs Mobile	r = 0.996	r = 0.994
Desktop vs Mobile2	r = 0.996	r = 0.994
Mobile1 vs Mobile2	r = 1.000	r = 1.000

5.5. Fundamental Algorithm Performance Analysis

5.5.1. Comprehensive Algorithm Assessment

Systematic evaluation of fundamental computer vision algorithms established performance hierarchies based on comprehensive statistical analysis across all datasets and algorithm combinations. The analysis reveals distinct performance characteristics supporting evidence-based algorithm selection:

AKAZE: Achieves highest detector performance (mean efficiency: 0.5472) with maximum performance of 0.8092 and moderate standard deviation (0.1771) across 179 evaluations, establishing superior consistency across operational scenarios. AKAZE's balanced performance characteristics make it optimal for applications requiring predictable results across diverse UAV deployment conditions.

SIFT: Demonstrates highest descriptor performance (mean efficiency: 0.6354) with maximum performance of 0.8092 across 129 evaluations, providing stable correspondence quality. SIFT's extensive evaluation coverage and consistent high-quality results confirm its established reputation for reliability in challenging imaging scenarios typical of UAV operations.

ORB: Achieves highest algorithm efficiency (mean efficiency: 0.7483) across all combinations with peak efficiency capability of 0.8821, demonstrating significant versatility when optimally configured. This characteristic makes ORB particularly suitable for applications where algorithm parameters can be systematically optimized for specific operational requirements.

5.5.2. Algorithm Versatility Assessment

Algorithm versatility analysis reveals distinct deployment characteristics supporting evidence-based selection for diverse UAV applications:

ORB: Highest average efficiency (0.7483) with substantial peak efficiency capability (0.8821), making it optimal for applications requiring both high efficiency and reliable consistency across diverse operational conditions. ORB demonstrates strong results across different deployment scenarios.

SIFT: Strong second-place performance (0.7254 mean efficiency) with reliable maximum performance (0.8198), supporting its continued utility in applications requiring stable correspondence quality. SIFT maintains consistent high-quality results across challenging imaging conditions characteristic of UAV operations.

AKAZE: Solid third-place performance (0.7102 mean efficiency) with good consistency (StdDev: 0.0688), indicating reliable performance for applications requiring predictable results when algorithm parameters need standardization across different operational requirements.

Efficiency Insights: The analysis reveals that highest average efficiency (ORB), most consistent efficiency (KAZE with StdDev: 0.0553), and highest peak efficiency (ORB) represent distinct algorithmic characteristics. These findings inform algorithm selection strategies based on specific operational requirements: ORB for high-efficiency applications, KAZE for consistency-critical deployments, and SIFT for balanced efficiency across diverse applications.

Table 5.11 Algorithm Component Performance Summary

Component Type	Category	Algorithm	Mean CUES	Peak CUES
Feature Detectors	Top Performer	AKAZE	0.5472	0.8092
Feature Detectors	Peak Performance	ORB	0.5297	0.8821
Feature Detectors	Balanced	STAR	0.5070	0.8383
Feature Descriptors	Top Performer	SIFT	0.6354	0.8092
Feature Descriptors	Second Best	DAISY	0.6213	0.8286
Feature Descriptors	Best Binary	BEBLID	0.5801	0.8489
Matching Strategies	Accuracy Priority	Brute Force	0.5234	0.8821
Matching Strategies	Speed Priority	FLANN	0.5187	0.8789
Distance Metrics	Floating-Point	L2	0.5456	0.8367
Distance Metrics	Binary	Hamming	0.4965	0.8821

Table 5.11 presents a consolidated summary of algorithm component performance, enabling direct comparison across all component types and providing definitive hierarchies for evidence-based algorithm selection.

This hierarchical analysis enables targeted algorithm selection: AKAZE provides the most consistent detector performance across diverse scenarios, while ORB achieves the highest peaks under optimal conditions. SIFT maintains descriptor superiority for accuracy-critical applications, while BEBLID offers a competitive binary alternative for resource-constrained deployments. Brute Force matching proves optimal for typical UAV feature densities, challenging conventional FLANN preferences.

5.6. Practical Deployment Implications

5.6.1. Evidence-Based Algorithm Selection

The comprehensive framework evaluation provides definitive guidance for algorithm selection across diverse UAV operational requirements:

Maximum Cross-Dataset Performance: The ORB-BEBLID-HAM-BF combination represents the optimal choice for diverse operational conditions, achieving top performance rankings across four of five datasets with efficiency scores ranging from 0.6219±0.0564 on Oxford to 0.8489±0.0095 on Drone (95% CI). The narrow confidence interval on the Drone dataset (±0.0095) indicates exceptional measurement stability, while this combination's consistency establishes its reliability for operational UAV deployments requiring robust performance across varying environmental conditions.

Controlled Environment Applications: The ORB-ORB-HAM-BF combination achieves peak efficiency (0.8821±0.0371, 95% CI) under synthetic conditions, making it optimal for applications with predictable imaging scenarios, indoor environments, or systematic parameter optimization capabilities. The confidence interval confirms statistically significant superiority over alternative combinations.

Consistent Performance Requirements: AKAZE-based combinations, particularly AKAZE-SIFT (mean efficiency 0.7059), provide reliable performance across evaluation scenarios with minimal variation. These combinations suit applications requiring predictable results independent of specific operational conditions.

Resource-Constrained Deployments: Binary descriptor combinations (BEBLID, TEBLID) demonstrate competitive performance while maintaining computational efficiency advantages crucial for mobile UAV platforms and real-time processing requirements.

5.6.2. Application-Specific Recommendations

Based on comprehensive statistical analysis across 2,282 valid evaluations with 95% confidence interval quantification, specific algorithm recommendations emerge for distinct UAV application categories:

High-Performance UAV 3D Reconstruction: ORB-BEBLID-HAM-BF provides optimal cross-dataset consistency (efficiency scores 0.6219±0.0564 on Oxford to 0.8489±0.0095 on Drone, 95% CI) suitable for professional mapping, surveying, and 3D reconstruction applications requiring reliable performance across diverse operational environments. The narrow confidence intervals, particularly on the Drone dataset (±0.0095), confirm high measurement precision and algorithm stability.

Controlled Environment Applications: ORB-ORB-HAM-BF achieves peak synthetic performance (0.8821±0.0371, 95% CI) optimal for indoor mapping, warehouse inventory, infrastructure inspection, or applications with controlled lighting and systematic imaging protocols. The statistical significance compared to second-ranked alternatives confirms its superiority in controlled conditions.

Precision Correspondence Requirements: AKAZE-SIFT combinations (mean efficiency 0.7059) maintain superior accuracy in challenging scenarios where correspondence quality outweighs computational efficiency, suitable for archaeological documentation, heritage preservation, or high-precision measurement applications.

Mobile and Edge Computing: Binary descriptor combinations (ORB-BEBLID, ORB-TEBLID) demonstrate consistent cross-platform performance enabling reliable operation across server and mobile hardware architectures without significant algorithm reconfiguration, suitable for autonomous UAV systems and real-time processing applications.

Operational Versatility: SIFT-based combinations provide balanced performance characteristics suitable for applications requiring adaptability across unknown operational conditions, emergency response scenarios, or research applications where operational parameters cannot be predicted or controlled.

5.7. Validation and Statistical Significance

5.7.1. Methodology Validation Results

The CUES methodology demonstrated strong statistical robustness across the comprehensive evaluation framework. Analysis of 2,282 valid measurements from 784 algorithm combinations across five datasets confirmed methodology reliability through consistent algorithmic rankings and cross-dataset performance correlations averaging 0.865. Bootstrap-based confidence interval estimation (95% CI) provides uncertainty quantification for individual method combinations, enabling statistical significance assessment between competing algorithms.

Cross-dataset ranking stability analysis revealed notable consistency across all pairwise dataset comparisons with correlation coefficients ranging from 0.738 to 0.967. The comprehensive stability assessment demonstrates that algorithmic rankings remain consistent across diverse operational scenarios, validating the methodology's capacity to identify genuine performance characteristics independent of dataset-specific artifacts.

Statistical significance testing confirmed that performance differences between top-tier algorithm combinations represent authentic algorithmic characteristics rather than measurement noise. For example, the difference between ORB-BEBLID-HAM-BF (0.8489±0.0095) and ORB-TEBLID-HAM-BF (0.8247±0.0096) on the Drone dataset is statistically significant as their 95% confidence intervals do not overlap, validating the performance hierarchy. The methodology successfully identified distinct performance hierarchies with confidence intervals supporting evidence-based algorithm selection decisions for UAV-based 3D reconstruction applications across diverse operational requirements and deployment constraints.

5.7.2. Performance Distribution Analysis

The comprehensive evaluation revealed significant performance variations across the 2,282 valid algorithm combinations with average cross-dataset performance of 0.7744 and performance spanning from 0.6219 to 0.8821. Performance standard deviation of 0.1032 across datasets indicates meaningful algorithmic differences that justify systematic evaluation approaches for UAV applications.

The analysis identified 421 mutually valid combinations across all datasets, enabling reliable comparative analysis supporting evidence-based algorithm selection for diverse UAV operational requirements. Statistical analysis confirmed that 17 algorithm combinations achieved consistent top-tier performance across multiple datasets, establishing reliable performance hierarchies with statistical confidence. Table 5.12 summarizes the dataset-specific performance characteristics.

Table 5.12 Dataset-Specific Performance Characteristics Summary

Characteristic	Value	Description
Best Dataset Performance	0.8821±0.0371	ORB-ORB-HAM-BF on Synthetic dataset (95% CI)
Lowest Dataset Performance	0.6219±0.0564	ORB-BEBLID-HAM-BF on Oxford dataset (95% CI)
Cross-Dataset Stability	0.865	Average ranking correlation across datasets
Performance Variation	0.2602	Difference between best and lowest scores (0.8821 - 0.6219)

These distributions validate the methodology's capacity to distinguish meaningful performance differences while maintaining statistical rigor across diverse UAV operational scenarios, supporting confident algorithm selection decisions based on empirical evidence rather than subjective assessment.

5.7.3. Parameter Sensitivity Validation

The methodological robustness of CUES depends fundamentally on two parameter choices that define the evaluation framework: the normalization function exponent alpha and the weight aggregation strategy. While Section 3.2.6 provided theoretical justification for these selections grounded in multi-criteria decision making literature, empirical validation across operational datasets remains essential to demonstrate that algorithmic rankings remain stable under parameter variations. Sensitivity analysis addresses whether alternative parameterizations would substantively alter the conclusions derived from this evaluation framework, thereby establishing confidence bounds on the methodology's reliability for practical deployment decisions.

This validation employed systematic parameter perturbation across all five evaluation datasets. For the normalization parameter, alpha values spanning 0.1 to 0.75 were evaluated to assess ranking stability across the spectrum from aggressive compression to near-linear scaling. For weight aggregation, four mathematically distinct combination strategies were compared to determine whether the arithmetic mean approach introduces systematic bias relative to alternative aggregation methods. The analysis quantifies ranking consistency through Spearman rank correlation, providing direct measurement of how algorithmic hierarchies respond to parameter modifications.

Normalization Parameter Sensitivity

The nonlinear normalization function x_norm = ((x - min) / (max - min))^alpha transforms raw metric values into a standardized range while controlling sensitivity to performance outliers. The baseline alpha=0.5 square root transformation represents a balanced choice between linear preservation of relationships at higher alpha values and aggressive compression at lower values. To validate this selection, complete CUES rankings were recalculated for alpha values of 0.1, 0.25, 0.5, and 0.75, generating independent algorithmic hierarchies for each parameterization.

Table 5.13 presents the complete Spearman rank correlation matrix for the Drone dataset, demonstrating ranking stability across the full alpha parameter range. This comprehensive analysis reveals how algorithmic hierarchies respond to systematic parameter variations.

Detailed rank change analysis for the Drone dataset reveals exceptional stability: when comparing alpha=0.5 to alpha=0.25, all 10 top algorithms maintained stable rankings (±1 position change), with zero moderate or large rank perturbations. Similarly, comparison with alpha=0.75 showed 10 stable algorithms with no significant rank changes. Even the more extreme comparison with alpha=0.10 demonstrated 9 stable algorithms with only 1 moderate change (±2-3 positions) and zero large perturbations.

Table 5.13 Alpha Parameter Sensitivity Correlation Matrix (Drone)

alpha	0.10	0.25	0.50	0.75
0.10	1.0000	0.9955	0.9841	0.9725
0.25	0.9955	1.0000	0.9954	0.9865
0.50	0.9841	0.9954	1.0000	0.9967
0.75	0.9725	0.9865	0.9967	1.0000

Table 5.14 summarizes the Spearman rank correlation between the baseline alpha=0.5 configuration and its adjacent parameter values at alpha=0.25 and alpha=0.75 across all datasets. The minimum correlation represents the most conservative stability estimate, capturing the maximum ranking perturbation that occurs when alpha deviates by 50 percent from the baseline value.

The consistently high correlations, averaging ρ=0.9968 across all datasets, demonstrate that algorithmic rankings remain highly stable across substantial variations in the normalization parameter. The Drone dataset exhibits the lowest correlation at ρ=0.9954, yet even this represents only minor rank perturbations concentrated among algorithmically similar methods. The UAV and Synthetic datasets show particularly strong stability with ρ exceeding 0.9975, indicating near-invariance to parameter selection. This empirical stability validates the theoretical justification for alpha=0.5 while confirming that conclusions regarding optimal algorithm combinations remain reliable across the reasonable parameter space.

Table 5.14 Normalization Parameter Sensitivity Analysis Across All Datasets

Dataset	rho (alpha=0.5 vs 0.25)	rho (alpha=0.5 vs 0.75)	Minimum rho
Drone	0.9954	0.9967	0.9954
UAV	0.9976	0.9988	0.9976
AirSim	0.9977	0.9965	0.9965
Oxford	0.9970	0.9988	0.9970
Synthetic	0.9977	0.9984	0.9977
Average	0.9971	0.9978	0.9968

Weight Aggregation Method Sensitivity

The CUES methodology combines four unsupervised weighting approaches through arithmetic mean aggregation, treating entropy, CRITIC, PCA, and variance weights with equal influence. This equal-treatment principle is theoretically justified because each weighting method captures distinct statistical properties: entropy measures information content and discriminative power, PCA identifies variance-explained dimensionality importance, CRITIC combines contrast intensity with intercriteria correlation analysis, and variance directly measures distribution breadth.

The arithmetic mean provides a robust combination when weights operate on similar scales [0,1], following common practice in ensemble weighting methods.

Four mathematically distinct aggregation methods were evaluated as alternatives to the arithmetic mean baseline. The geometric mean emphasizes proportional relationships and reduces sensitivity to extreme weight values through multiplicative combination. The harmonic mean provides conservative weighting that emphasizes lower-magnitude components, reducing influence of potentially unstable high weights. The trimmed mean with 10 percent trim discards extreme weight values, providing robustness against outliers in the weight distributions.

Table 5.15 presents the complete Spearman rank correlation matrix between aggregation methods for the Drone dataset, demonstrating the remarkable consistency of algorithmic rankings regardless of mathematical aggregation approach.

Table 5.15 Drone Weight Aggregation Sensitivity Correlation Matrix

Method	Arithmetic	Geometric	Harmonic	Trimmed
Arithmetic	1.0000	0.9999	0.9997	0.9992
Geometric	0.9999	1.0000	0.9999	0.9995
Harmonic	0.9997	0.9999	1.0000	0.9997
Trimmed	0.9992	0.9995	0.9997	1.0000

Table 5.16 presents the average Spearman rank correlation between arithmetic mean rankings and each alternative aggregation method, computed across all five datasets. These correlations quantify whether alternative mathematical formulations of weight combination produce substantially different algorithmic hierarchies.

Table 5.16 Weight Aggregation Method Sensitivity Analysis

Comparison	Average rho (across datasets)
Arithmetic vs Geometric	0.9998
Arithmetic vs Harmonic	0.9995
Arithmetic vs Trimmed	0.9993
Overall Average	0.9996

The correlations demonstrate near-perfect ranking stability, with the overall average of rho=0.9996 indicating that algorithmic hierarchies remain essentially invariant to aggregation method selection. The geometric mean shows the strongest agreement at rho=0.9998, differing only in cases where extreme weight ratios distinguish otherwise similar algorithms. Even the trimmed mean, which fundamentally alters the weight distribution by discarding extreme values, maintains rho=0.9993 correlation with the baseline approach.

To examine how aggregation methods differ in their treatment of individual metrics, Table 5.17 presents the normalized weight vectors for each aggregation approach using the Drone dataset as a representative example. These weights reflect the relative importance assigned to each evaluation metric after aggregation.

The weight distributions show minor variations across methods, with all approaches emphasizing quality metrics (Precision, Recall, Repeatability, F1 Score contributing 0.5476-0.5964 combined weight) over timing metrics (Total Time, Inlier Time contributing only 0.0019-0.0113 combined weight). This consistent emphasis on quality metrics across all aggregation methods confirms that the CUES methodology appropriately prioritizes correspondence quality while maintaining computational efficiency as a secondary consideration.

Table 5.17 Metric Weights by Aggregation Method (Drone Dataset)

Metric	Arithmetic	Geometric	Harmonic	Trimmed
Precision	0.1597	0.1643	0.1669	0.1724
Recall	0.1271	0.1347	0.1412	0.1375
Repeatability	0.1250	0.1326	0.1390	0.1344
F1 Score	0.1358	0.1430	0.1489	0.1485
Inliers	0.1098	0.1145	0.1182	0.1107
Matches	0.0776	0.0764	0.0755	0.0671
Total Time	0.0013	0.0002	0.0000	0.0006
Inlier Time	0.0100	0.0051	0.0019	0.0093

Sensitivity Analysis Conclusions

The dual sensitivity analysis confirms that CUES rankings are highly stable across parameter variations, with correlations consistently exceeding rho=0.99 across both dimensions. The alpha parameter sensitivity analysis demonstrates that rankings remain stable with average correlation rho=0.9968, validating the choice of alpha=0.5 (square root transformation) as appropriate. The weight aggregation method sensitivity analysis shows even stronger stability with average correlation rho=0.9996, confirming that the arithmetic mean approach is well justified.

These findings are supported by four key factors: (1) high rank stability across alpha values confirms the square root transformation is appropriate, (2) high rank stability across aggregation methods validates the arithmetic mean combination, (3) theoretical foundations in multi-criteria decision making literature provide principled justification, and (4) consistent results across all five evaluation datasets demonstrate generalizability. This comprehensive validation establishes confidence bounds on the methodology's reliability, supporting evidence-based algorithm selection decisions for practical UAV deployment scenarios.

6. CONCLUSION

This research successfully developed and validated a comprehensive objective evaluation methodology for UAV 3D reconstruction algorithm selection, addressing the critical gap between theoretical algorithm capabilities and practical deployment decisions. The systematic evaluation of 784 algorithm combinations across five diverse datasets produced 2,282 valid measurements, establishing the first comprehensive evidence-based framework for UAV-based 3D reconstruction applications with statistical rigor and cross-platform validation.

6.1. Methodological Contributions

The CUES represents the primary methodological innovation, integrating four distinct unsupervised weighting approaches through principled mathematical combination. Entropy weighting uses information theory principles for discriminative metric emphasis. CRITIC weighting combines contrast intensity measurement with intercriteria correlation analysis. PCA weighting employs dimensionality reduction to identify variance-maximizing evaluation criteria. Variance weighting directly measures discriminative power through performance distribution analysis.

Comprehensive statistical validation demonstrates solid methodology robustness through cross-dataset ranking correlations averaging 0.865, with individual correlations ranging from 0.738 to 0.967 across all dataset pairs. Bootstrap-based 95% confidence interval estimation enables statistical significance assessment between algorithm combinations, providing uncertainty quantification essential for evidence-based selection decisions. This consistency confirms the methodology's capacity to identify genuine algorithmic characteristics independent of dataset-specific artifacts, establishing reliable foundations for evidence-based algorithm selection in UAV applications.

6.2. Empirical Findings

6.2.1. Algorithm Performance Hierarchy

The comprehensive evaluation reveals definitive performance hierarchies across 2,282 valid evaluations with 95% confidence interval quantification. The ORB-BEBLID-HAM-BF combination emerges as the most consistent cross-dataset performer, achieving optimal rankings across four of five datasets with scores ranging from 0.6219 ± 0.0564 on Oxford to 0.8489 ± 0.0095 on Drone. The narrow confidence interval on the Drone dataset confirms exceptional measurement precision and algorithm stability.

Component analysis establishes clear algorithmic rankings: AKAZE achieves highest mean detector efficiency (0.5472), while SIFT demonstrates superior descriptor capabilities. ORB exhibits substantial overall algorithm efficiency (0.7483 mean efficiency) with the highest peak efficiency potential (0.8821±0.0371, 95% CI on Synthetic), indicating superior optimization potential across diverse operational conditions.

The AKAZE-SIFT detector-descriptor combination achieves optimal pairing performance (0.7059 mean) across 10 evaluation variants, establishing it as the premier combination for applications requiring consistent high-quality correspondence establishment. Binary descriptors, particularly BEBLID and TEBLID, demonstrate competitive performance while providing computational advantages crucial for resource-constrained UAV operations.

6.2.2. Cross-Platform Validation

Cross-platform analysis across Desktop, Mobile1 (Snapdragon 855), and Mobile2 (Dimensity 7200 Ultra) platforms reveals strong ranking consistency with correlation coefficients exceeding 0.994 for both Oxford and Synthetic datasets. The ORB-BEBLID-HAM-BF combination maintains optimal performance across all computational architectures with minimal performance variation (±0.02 points).

Statistical analysis identifies 9 algorithms maintaining top-10 rankings across all environments for Oxford dataset and 8 algorithms for Synthetic dataset, demonstrating robust algorithmic characteristics independent of computational architecture. Mobile platforms achieve competitive performance while maintaining timing-to-quality ratios between 0.074-0.096, confirming the feasibility of high-performance algorithm deployment on resource-constrained UAV hardware.

6.3. Practical Applications

The comprehensive evaluation provides definitive operational guidance for diverse UAV application scenarios with 95% confidence interval quantification supporting statistically rigorous selection decisions. For maximum cross-dataset performance, ORB-BEBLID-HAM-BF represents the optimal choice, demonstrating consistent performance across operational environments ranging from controlled synthetic conditions (0.825±0.062) to challenging real-world scenarios (0.849±0.010 on Drone).

Controlled environment applications achieve peak efficiency with ORB-ORB-HAM-BF (0.8821±0.0371, 95% CI), while precision correspondence requirements benefit from AKAZE-SIFT combinations (0.7059 mean). Resource-constrained mobile deployments favor binary descriptor combinations that maintain competitive efficiency while enabling reliable cross-platform operation without algorithm reconfiguration.

The research identifies 421 mutually valid combinations across all datasets, enabling reliable comparative analysis that informs evidence-based algorithm selection for UAV applications with specific performance requirements or deployment constraints. Performance variations spanning from

0.6219 to 0.8821 between optimal and suboptimal selections demonstrate the critical importance of systematic algorithm evaluation for operational UAV systems.

6.4. Research Impact and Validation

Statistical validation establishes methodology reliability through comprehensive cross-dataset consistency analysis spanning five diverse datasets with correlation coefficients ranging from 0.738 to 0.967. Key performance differences between leading algorithm combinations achieve statistical significance, confirming that the identified hierarchies represent genuine algorithmic characteristics rather than measurement artifacts.

The research contributes the first comprehensive objective evaluation methodology for UAV-based 3D reconstruction algorithm selection, establishing new standards for algorithm evaluation in computer vision applications. Open-source availability of all materials, including 2,282 validated measurements, interactive visualization tools, and comprehensive performance databases, promotes collaborative advancement while ensuring scientific reproducibility and enabling the research community to build upon these empirical foundations.

6.5. Limitations and Future Directions

The evaluation concentrates on traditional OpenCV algorithms without systematic integration of contemporary deep learning approaches that increasingly influence computer vision research. While this focus provides essential baseline understanding and addresses practical deployment requirements for resource-constrained UAV systems, future extensions should incorporate neural network-based methods such as SuperPoint, D2-Net, and LoFTR within the objective evaluation framework.

Dataset scope encompasses RGB imagery across diverse operational scenarios but lacks systematic evaluation of multispectral, hyperspectral, or thermal data streams increasingly important in specialized UAV applications. Hardware evaluation emphasizes CPU-based processing without comprehensive assessment of GPU acceleration, specialized vision processors, or emerging edge computing architectures that may significantly alter performance characteristics.

Future research directions include methodological extensions to video-based reconstruction evaluation, multi-sensor fusion assessment, and real-time performance evaluation incorporating frame rate constraints and latency requirements. Application-specific adaptations for precision agriculture, infrastructure inspection, emergency response, and autonomous navigation represent important directions that could extend the practical utility and operational relevance of the evaluation framework while maintaining the objective, statistically rigorous foundations established in this research.

6.6. Final Assessment

This thesis establishes new standards for objective algorithm evaluation in UAV-based 3D reconstruction through theoretically grounded, statistically validated methodologies that demonstrate exceptional cross-dataset consistency and cross-platform reliability. The systematic evaluation of 784 algorithm combinations producing 2,282 valid measurements provides the research community with unprecedented empirical foundations for evidence-based algorithm selection decisions in operational UAV applications.

The CUES methodology contributes a reusable, adaptable framework extending beyond UAV applications to other computer vision evaluation challenges requiring objective, multi-criteria assessment. The comprehensive statistical validation with cross-dataset correlations averaging 0.865 establishes confidence in the methodology's capacity to identify genuine algorithmic characteristics independent of specific operational conditions.

The commitment to open science principles through comprehensive data sharing, interactive visualization tools, and validated performance databases ensures that the research community can build upon these foundations to drive continued advancement in efficient, accurate, and reliable UAV-based 3D reconstruction capabilities. This research provides immediate practical value while establishing methodological foundations that will support future innovations in UAV-based 3D reconstruction and broader computer vision applications requiring objective algorithm evaluation with statistical rigor.

REFERENCES

Abeho, D. R., Shoko, M., & Odera, P. A. (2024). Effects of camera calibration on the accuracy of Unmanned Aerial Vehicle sensor products. *International Journal of Engineering and Geosciences*. https://doi.org/10.26833/ijeg.1422619

Agrawal, M., Konolige, K., & Blas, M. R. (2008). CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching. In D. A. Forsyth, P. H. S. Torr, & A. Zisserman (Eds.), *Computer Vision - ECCV 2008, 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part IV* (Vol. 5305, pp. 102–115). Springer. https://doi.org/10.1007/978-3-540-88693-8_8

Alahi, A., Ortiz, R., & Vandergheynst, P. (2012). FREAK: Fast retina keypoint. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 510–517. https://doi.org/10.1109/CVPR.2012.6247715

Alcantarilla, P. F., Bartoli, A., & Davison, A. J. (2012). KAZE Features. In A. W. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI* (Vol. 7577, pp. 214–227). Springer. https://doi.org/10.1007/978-3-642-33783-3_16

Alcantarilla, P. F., Nuevo, J., & Bartoli, A. (2013). Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces. In T. Burghardt, D. Damen, W. W. Mayol-Cuevas, & M. Mirmehdi (Eds.), *British Machine Vision Conference, BMVC 2013, Bristol, UK, September 9-13, 2013*. BMVA Press. https://doi.org/10.5244/C.27.13

Amaral, L. R. do, Zerbato, C., Freitas, R. G. de, Barbosa Júnior, M. R., & Simões, I. O. P. da S. (2020). UAV applications in Agriculture 4.0. *REVISTA CIÊNCIA AGRONÔMICA*, *51*(5). https://doi.org/10.5935/1806-6690.20200091

Arnaldo, C. G., Moreno, F. P., Suárez, M. Z., & Jurado, R. D.-A. (2025). Autonomous Route Planning for UAV-Based 3D Reconstruction. *EASN 2024*, 78. https://doi.org/10.3390/engproc2025090078

Balntas, V., Lenc, K., Vedaldi, A., & Mikolajczyk, K. (2017). *HPatches: A benchmark and evaluation of handcrafted and learned local descriptors*. http://arxiv.org/abs/1704.05939

Baráth, D., Noskova, J., Ivashechkin, M., & Matas, J. (2020). MAGSAC++, a Fast, Reliable and Accurate Robust Estimator. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, 1301–1309. https://doi.org/10.1109/CVPR42600.2020.00138

Barba, S., Barbarella, M., Di Benedetto, A., Fiani, M., Gujski, L., & Limongiello, M. (2019). Accuracy Assessment of 3D Photogrammetric Models from an Unmanned Aerial Vehicle. *Drones*, *3*(4), 79. https://doi.org/10.3390/drones3040079

Bazi, Y., Malek, S., Alajlan, N., & Alhichri, H. (2014). An automatic approach for palm tree counting in UAV images. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 537–540. https://doi.org/10.1109/IGARSS.2014.6946478

Bianco, S., Ciocca, G., & Marelli, D. (2018). Evaluating the performance of structure from motion pipelines. *Journal of Imaging*, *4*(8). https://doi.org/10.3390/jimaging4080098

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). *BRIEF: Binary Robust Independent Elementary Features*.

Cao, M., Jia, W., Li, S., Li, Y., Zheng, L., & Liu, X. (2019). GPU-accelerated feature tracking for 3D reconstruction. *Optics and Laser Technology*, *110*, 165–175. https://doi.org/10.1016/j.optlastec.2018.08.045

Cao, M., Jia, W., Lv, Z., Li, Y., Xie, W., Zheng, L., & Liu, X. (2019). Fast and robust feature tracking for 3D reconstruction. *Optics and Laser Technology*, *110*, 120–128. https://doi.org/10.1016/j.optlastec.2018.05.036

Chen, C., Wang, B., Lu, C. X., Trigoni, N., & Markham, A. (2023). *Deep Learning for Visual Localization and Mapping: A Survey*. http://arxiv.org/abs/2308.14039

Clapuyt, F., Vanacker, V., & Van Oost, K. (2016). Reproducibility of UAV-based earth topography reconstructions based on Structure-from-Motion algorithms. *Geomorphology*, *260*, 4–15. https://doi.org/10.1016/j.geomorph.2015.05.011

Crommelinck, S., Bennett, R., Gerke, M., Nex, F., Yang, M. Y., & Vosselman, G. (2016). Review of automatic feature extraction from high-resolution optical sensor data for UAV-based cadastral mapping. In *Remote Sensing* (Vol. 8, Number 8). MDPI AG. https://doi.org/10.3390/rs8080689

DeTone, D., Malisiewicz, T., & Rabinovich, A. (2018). SuperPoint: Self-Supervised Interest Point Detection and Description. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 337–33712. https://doi.org/10.1109/CVPRW.2018.00060

Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Comput. Oper. Res.*, *22*(7), 763–770. https://doi.org/10.1016/0305-0548(94)00059-H

Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A., & Sattler, T. (2019). D2-Net: A Trainable CNN for Joint Description and Detection of Local Features. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8084–8093. https://doi.org/10.1109/CVPR.2019.00828

Fernández Alcantarilla, P., Bartoli, A., & Davison, A. J. (2012). *KAZE Features*.

Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Flores-de-Santiago, F., Valderrama-Landeros, L., Rodríguez-Sobreyra, R., & Flores-Verdugo, F. (2020). Assessing the effect of flight altitude and overlap on orthoimage generation for UAV estimates of coastal wetlands. *Journal of Coastal Conservation*, *24*(3), 35. https://doi.org/10.1007/s11852-020-00753-9

Gaffey, C., & Bhardwaj, A. (2020). Applications of unmanned aerial vehicles in cryosphere: Latest advances and prospects. *Remote Sensing*, *12*(6). https://doi.org/10.3390/rs12060948

Harris, C. G., & Stephens, M. (1988). A Combined Corner and Edge Detector. In C. J. Taylor (Ed.), *Proceedings of the Alvey Vision Conference, AVC 1988, Manchester, UK, September, 1988* (pp. 1–6). Alvey Vision Club. https://doi.org/10.5244/C.2.23

Hartley, R., & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811685

Jianbo Shi, & Tomasi. (1994). Good features to track. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, 593–600. https://doi.org/10.1109/CVPR.1994.323794

Jiang, S., Jiang, C., & Jiang, W. (2020). Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS Journal of Photogrammetry and Remote Sensing*, *167*, 230–251. https://doi.org/10.1016/j.isprsjprs.2020.04.016

Jiang, S., Jiang, W., Huang, W., & Yang, L. (2017). UAV-based oblique photogrammetry for outdoor data acquisition and offsite visual inspection of transmission line. *Remote Sensing*, *9*(3). https://doi.org/10.3390/rs9030278

Jiang, S., Jiang, W., & Wang, L. (2022). Unmanned Aerial Vehicle-Based Photogrammetric 3D Mapping: A survey of techniques, applications, and challenges. *IEEE Geoscience and Remote Sensing Magazine*, *10*(2), 135–171. https://doi.org/10.1109/MGRS.2021.3122248

Joshi, K., & Patel, M. I. (2020). Recent advances in local feature detector and descriptor: a literature survey. In *International Journal of Multimedia Information Retrieval* (Vol. 9, Number 4, pp. 231–247). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s13735-020-00200-3

Kedzierski, M., & Wierzbicki, D. (2015). Radiometric quality assessment of images acquired by UAV's in various lighting and weather conditions. *Measurement*, *76*, 156–169. https://doi.org/10.1016/j.measurement.2015.08.003

Kholil, M., Ismanto, I., & Fu'ad, M. N. (2021). 3D reconstruction using Structure From Motion (SFM) algorithm and Multi View Stereo (MVS) based on computer vision. *IOP Conference Series: Materials Science and Engineering*, *1073*(1), 012066. https://doi.org/10.1088/1757-899X/1073/1/012066

Lee, J., & Yoo, S. (2025). *Dense-SfM: Structure from Motion with Dense Consistent Matching*. https://arxiv.org/abs/2501.14277

Leem, J., Kim, J., Kang, I.-S., Choi, J., Song, J.-J., Mehrishal, S., & Shao, Y. (2025). Practical error prediction in UAV imagery-based 3D reconstruction: assessing the impact of image quality factors. *International Journal of Remote Sensing*, *46*(3), 1000–1030. https://doi.org/10.1080/01431161.2025.2450561

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011). BRISK: Binary Robust invariant scalable keypoints. *Proceedings of the IEEE International Conference on Computer Vision*, 2548–2555. https://doi.org/10.1109/ICCV.2011.6126542

Levi, G., & Hassner, T. (2016, May 23). LATCH: Learned arrangements of three patch codes. *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016*. https://doi.org/10.1109/WACV.2016.7477723

Li, X., Wang, K., Liu, L., Xin, J., Yang, H., & Gao, C. (2011). Application of the Entropy Weight and TOPSIS Method in Safety Evaluation of Coal Mines. *Procedia Engineering*, *26*, 2085–2091. https://doi.org/10.1016/j.proeng.2011.11.2410

Liao, Y., Di, Y., Zhu, K., Zhou, H., Lu, M., Zhang, Y., Duan, Q., & Liu, J. (2024). Local feature matching from detector-based to detector-free: a survey. *Applied Intelligence*, *54*(5), 3954–3989. https://doi.org/10.1007/s10489-024-05330-3

Liénard, J., Vogs, A., Gatziolis, D., & Strigul, N. (2016). Embedded, real-time UAV control for improved, image-based 3D scene reconstruction. *Measurement: Journal of the International Measurement Confederation*, *81*, 264–269. https://doi.org/10.1016/j.measurement.2015.12.014

Liu, L., Wang, C., Feng, Chuncheng, Gong, W., Zhang, L., Liao, L., & Feng, Chang. (2024). Incremental SFM 3D Reconstruction Based on Deep Learning. *Electronics*, *13*(14), 2850. https://doi.org/10.3390/electronics13142850

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. In *International Journal of Computer Vision* (Vol. 60, Number 2).

Luo, H., Zhang, J., Liu, X., Zhang, L., & Liu, J. (2024). Large-Scale 3D Reconstruction from Multi-View Imagery: A Comprehensive Review. *Remote Sensing*, *16*(5), 773. https://doi.org/10.3390/rs16050773

Ma, J., Jiang, X., Fan, A., Jiang, J., & Yan, J. (2021). Image Matching from Handcrafted to Deep Features: A Survey. *International Journal of Computer Vision*, *129*(1), 23–79. https://doi.org/10.1007/s11263-020-01359-2

Maes, W. H. (2025). Practical Guidelines for Performing UAV Mapping Flights with Snapshot Sensors. *Remote Sensing*, *17*(4), 606. https://doi.org/10.3390/rs17040606

Mair, E., Hager, G. D., Burschka, D., Suppa, M., & Hirzinger, G. (2010). *Adaptive and Generic Corner Detection Based on the Accelerated Segment Test* (pp. 183–196). https://doi.org/10.1007/978-3-642-15552-9_14

Man, K., & Chahl, J. (2022). A Review of Synthetic Image Data and Its Use in Computer Vision. *Journal of Imaging*, *8*(11). https://doi.org/10.3390/jimaging8110310

Marelli, D., Bianco, S., & Ciocca, G. (2020). IVL-SYNTHSFM-v2: A synthetic dataset with exact ground truth for the evaluation of 3D reconstruction pipelines. *Data in Brief*, *29*, 105041. https://doi.org/10.1016/j.dib.2019.105041

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, *22*(10 SPEC. ISS.), 761–767. https://doi.org/10.1016/j.imavis.2004.02.006

Medvedev, A., Telnova, N., Alekseenko, N., Koshkarev, A., Kuznetchenko, P., Asmaryan, S., & Narykov, A. (2020). UAV-Derived Data Application for Environmental Monitoring of the Coastal Area of Lake Sevan, Armenia with a Changing Water Level. *Remote Sensing*, *12*(22), 3821. https://doi.org/10.3390/rs12223821

Mendu, B., & Mbuli, N. (2025). State-of-the-Art Review on the Application of Unmanned Aerial Vehicles (UAVs) in Power Line Inspections: Current Innovations, Trends, and Future Prospects. *Drones*, *9*(4), 265. https://doi.org/10.3390/drones9040265

Mikolajczyk, K., & Schmid, C. (2004). Scale & Affine Invariant Interest Point Detectors. In *International Journal of Computer Vision* (Vol. 60, Number 1).

Mohsan, S. A. H., Othman, N. Q. H., Khan, M. A., Amjad, H., & Żywiołek, J. (2022). A Comprehensive Review of Micro UAV Charging Techniques. *Micromachines*, *13*(6), 977. https://doi.org/10.3390/mi13060977

Muhmad Kamarulzaman, A. M., Wan Mohd Jaafar, W. S., Mohd Said, M. N., Saad, S. N. M., & Mohan, M. (2023). UAV Implementations in Urban Planning and Related Sectors of Rapidly Developing Nations: A Review and Future Perspectives for Malaysia. *Remote Sensing*, *15*(11), 2845. https://doi.org/10.3390/rs15112845

Muja, M., & Lowe, D. G. (2009). Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In A. Ranchordas & H. Araújo (Eds.), *VISAPP 2009 - Proceedings of the Fourth International Conference on Computer Vision Theory and Applications, Lisboa, Portugal, February 5-8, 2009 - Volume 1* (pp. 331–340). INSTICC Press.

Ndlovu, H. S., Odindi, J., Sibanda, M., & Mutanga, O. (2024). A systematic review on the application of UAV-based thermal remote sensing for assessing and monitoring crop water status in crop farming systems. *International Journal of Remote Sensing*, *45*(15), 4923–4960. https://doi.org/10.1080/01431161.2024.2368933

Nesbit, P. R., & Hugenholtz, C. H. (2019). Enhancing UAV-SfM 3D model accuracy in high-relief landscapes by incorporating oblique images. *Remote Sensing*, *11*(3). https://doi.org/10.3390/rs11030239

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: a review. *Applied Geomatics*, *6*(1), 1–15. https://doi.org/10.1007/s12518-013-0120-x

Ozyesil, O., Voroninski, V., Basri, R., & Singer, A. (2017). *A Survey of Structure from Motion*. http://arxiv.org/abs/1701.08493

Partama, I. G. Y. (2025). 3D MODELING USING UAV-PHOTOGRAMMETRY TECHNIQUE FOR DIGITAL DOCUMENTATION OF CULTURAL HERITAGE BUILDINGS. *International Journal of GEOMATE*, *28*(126). https://doi.org/10.21660/2025.126.4768

Radoglou-Grammatikis, P., Sarigiannidis, P., Lagkas, T., & Moscholios, I. (2020). A compilation of UAV applications for precision agriculture. *Computer Networks, 172*. https://doi.org/10.1016/j.comnet.2020.107148

Richter, S. R., Vineet, V., Roth, S., & Koltun, V. (2016). Playing for Data: Ground Truth from Computer Games. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *European Conference on Computer Vision (ECCV)* (Vol. 9906, pp. 102–118). Springer International Publishing.

Rosten, E., & Drummond, T. (2006). *Machine Learning for High-Speed Corner Detection*.

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. *Proceedings of the IEEE International Conference on Computer Vision*, 2564–2571. https://doi.org/10.1109/ICCV.2011.6126544

Schichler, L., Festl, K., & Solmaz, S. (2025). Robust Multi-Sensor Fusion for Localization in Hazardous Environments Using Thermal, LiDAR, and GNSS Data. *Sensors*, *25*(7). https://doi.org/10.3390/s25072032

Schönberger, J. L., & Frahm, J.-M. (2016). Structure-from-Motion Revisited. *Conference on Computer Vision and Pattern Recognition (CVPR)*.

Schonberger, J. L., Hardmeier, H., Sattler, T., & Pollefeys, M. (2017). Comparative Evaluation of Hand-Crafted and Learned Local Features. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6959–6968. https://doi.org/10.1109/CVPR.2017.736

Seifert, E., Seifert, S., Vogt, H., Drew, D., van Aardt, J., Kunneke, A., & Seifert, T. (2019). Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. *Remote Sensing*, *11*(10). https://doi.org/10.3390/rs11101252

Shah, S., Dey, D., Lovett, C., & Kapoor, A. (2017). *AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles*. http://arxiv.org/abs/1705.05065

Shahbazi, M., Sohn, G., Théau, J., & Menard, P. (2015). Development and evaluation of a UAV-photogrammetry system for precise 3D environmental modeling. *Sensors (Switzerland), 15*(11), 27493–27524. https://doi.org/10.3390/s151127493

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Learning local feature descriptors using convex optimisation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(8), 1573–1585. https://doi.org/10.1109/TPAMI.2014.2301163

Song, Y. (2025). *Using synthetic data to improve the performance of UAV-based vehicle detection and speed estimation models*. http://essay.utwente.nl/106610/

Suárez, I., Buenaposada, J. M., & Baumela, L. (2021). *Revisiting Binary Local Image Description for Resource Limited Devices*. http://arxiv.org/abs/2108.08380

Suárez, I., Sfeir, G., Buenaposada, J. M., & Baumela, L. (2020). *BEBLID: Boosted Efficient Binary Local Image Descriptor*. www.elsevier.com

Sun, J., Shen, Z., Wang, Y., Bao, H., & Zhou, X. (2021). LoFTR: Detector-Free Local Feature Matching with Transformers. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8918–8927. https://doi.org/10.1109/CVPR46437.2021.00881

Taketomi, T., Uchiyama, H., & Ikeda, S. (2017). Visual SLAM algorithms: A survey from 2010 to 2016. In *IPSJ Transactions on Computer Vision and Applications* (Vol. 9). Springer. https://doi.org/10.1186/s41074-017-0027-2

Tola, E., Lepetit, V., & Fua, P. (2010). DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(5), 815–830. https://doi.org/10.1109/TPAMI.2009.77

Tombari, F., & Stefano, L. Di. (2014). Interest Points via Maximal Self-Dissimilarities. In D. Cremers, I. D. Reid, H. Saito, & M.-H. Yang (Eds.), *Computer Vision - ACCV 2014 - 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part II* (Vol. 9004, pp. 586–600). Springer. https://doi.org/10.1007/978-3-319-16808-1_39

Trzcinski, T., Christoudias, M., Fua, P., & Lepetit, V. (2013). Boosting binary keypoint descriptors. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2874–2881. https://doi.org/10.1109/CVPR.2013.370

Ulvi, A., & Hamal, S. N. G. (2025). Fusion of IPAD Pro LiDAR and SfM-Based Photogrammetry for 3D Documentation of Cultural Heritage. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*. https://doi.org/10.1007/s40996-025-01936-w

Vacca, G., Dessì, A., & Sacco, A. (2017). The use of nadir and oblique UAV images for building knowledge. *ISPRS International Journal of Geo-Information*, *6*(12). https://doi.org/10.3390/ijgi6120393

Vemprala, S., Chen, S., Shukla, A., Narayanan, D., & Kapoor, A. (2023). *GRID: A Platform for General Robot Intelligence Development*. https://arxiv.org/abs/2310.00887

Venkata Sai Chakradhar Reddy, D., Sahoo, R. N., Kondraju, T., Rejith, R. G., Ranjan, R., Bhandari, A., Moursy, A., Tripathi, S. C., & Kumar, N. (2025). Drone-Based Multispectral Imaging for Precision Monitoring of Crop Growth Variables. *IECAG 2024*, 10. https://doi.org/10.3390/blsf2025041010

Wang, J., Zhang, F., Zhang, Y., Liu, Y., & Cheng, T. (2023). Lightweight Object Detection Algorithm for UAV Aerial Imagery. *Sensors*, *23*(13), 5786. https://doi.org/10.3390/s23135786

Wang, T., Guan, T., Qiu, F., Liu, L., Zhang, X., Zeng, H., & Zhang, Q. (2025). Evaluation of Scale Effects on UAV-Based Hyperspectral Imaging for Remote Sensing of Vegetation. *Remote Sensing*, *17*(6), 1080. https://doi.org/10.3390/rs17061080

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). "Structure-from-Motion" photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, *2*(1–3), 37–52.

Xiang, T.-Z., Xia, G.-S., & Zhang, L. (2019). *Mini-Unmanned Aerial Vehicle-Based Remote Sensing: Techniques, Applications, and Prospects*.

Xu, Y., Monasse, P., Geraud, T., & Najman, L. (2014). Tree-based morse regions: A topological approach to local feature detection. *IEEE Transactions on Image Processing*, *23*(12), 5612–5625. https://doi.org/10.1109/TIP.2014.2364127

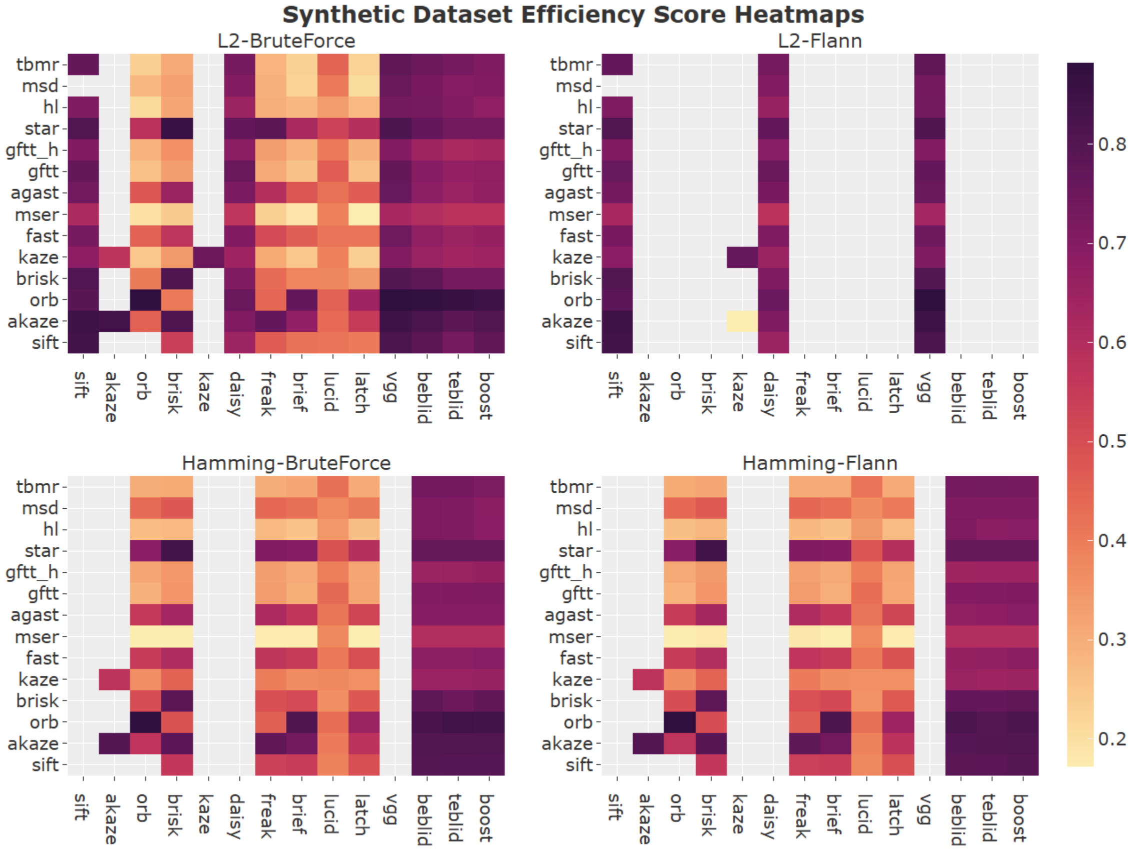
Ziegler, A., Christiansen, E. M., Kriegman, D. J., & Belongie, S. J. (2012). Locally Uniform Comparison Image Descriptor. In P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States* (pp. 1–9). https://proceedings.neurips.cc/paper/2012/hash/c20ad4d76fe97759aa27a0c99bff6710-Abstract.html

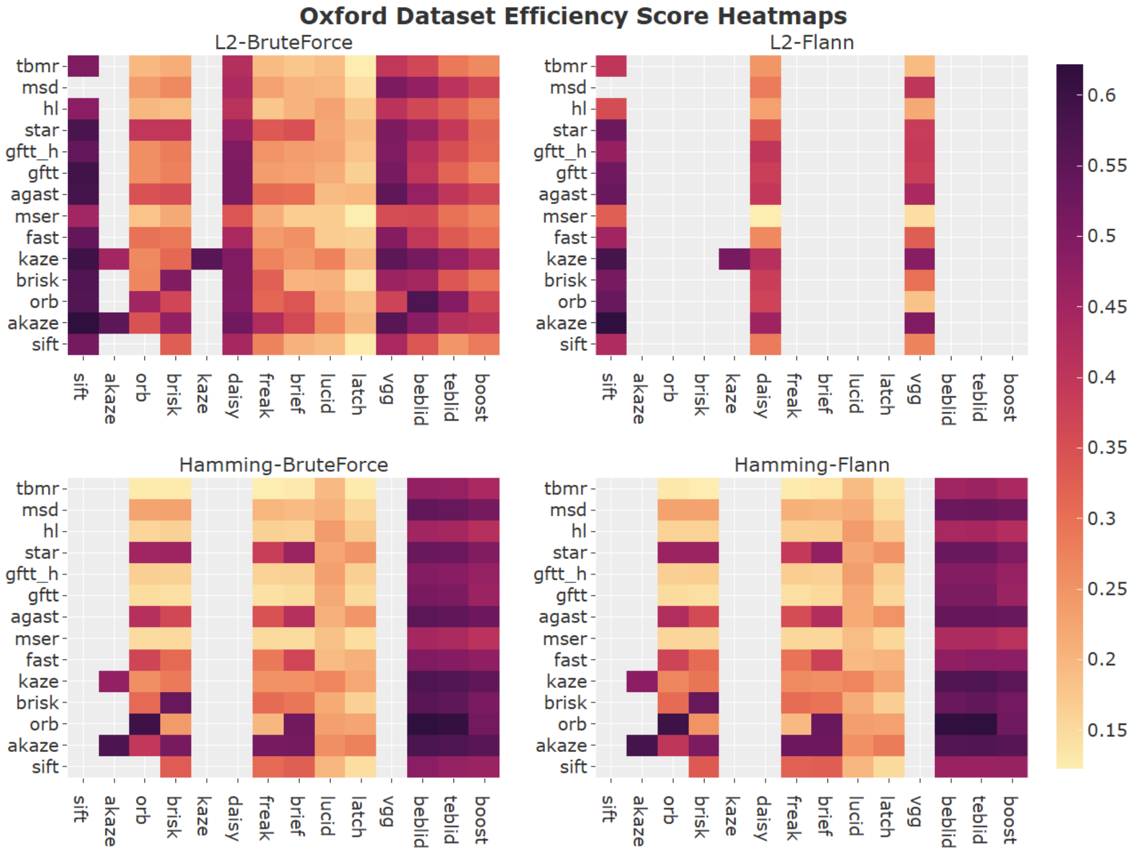
CURRICULUM VITAE

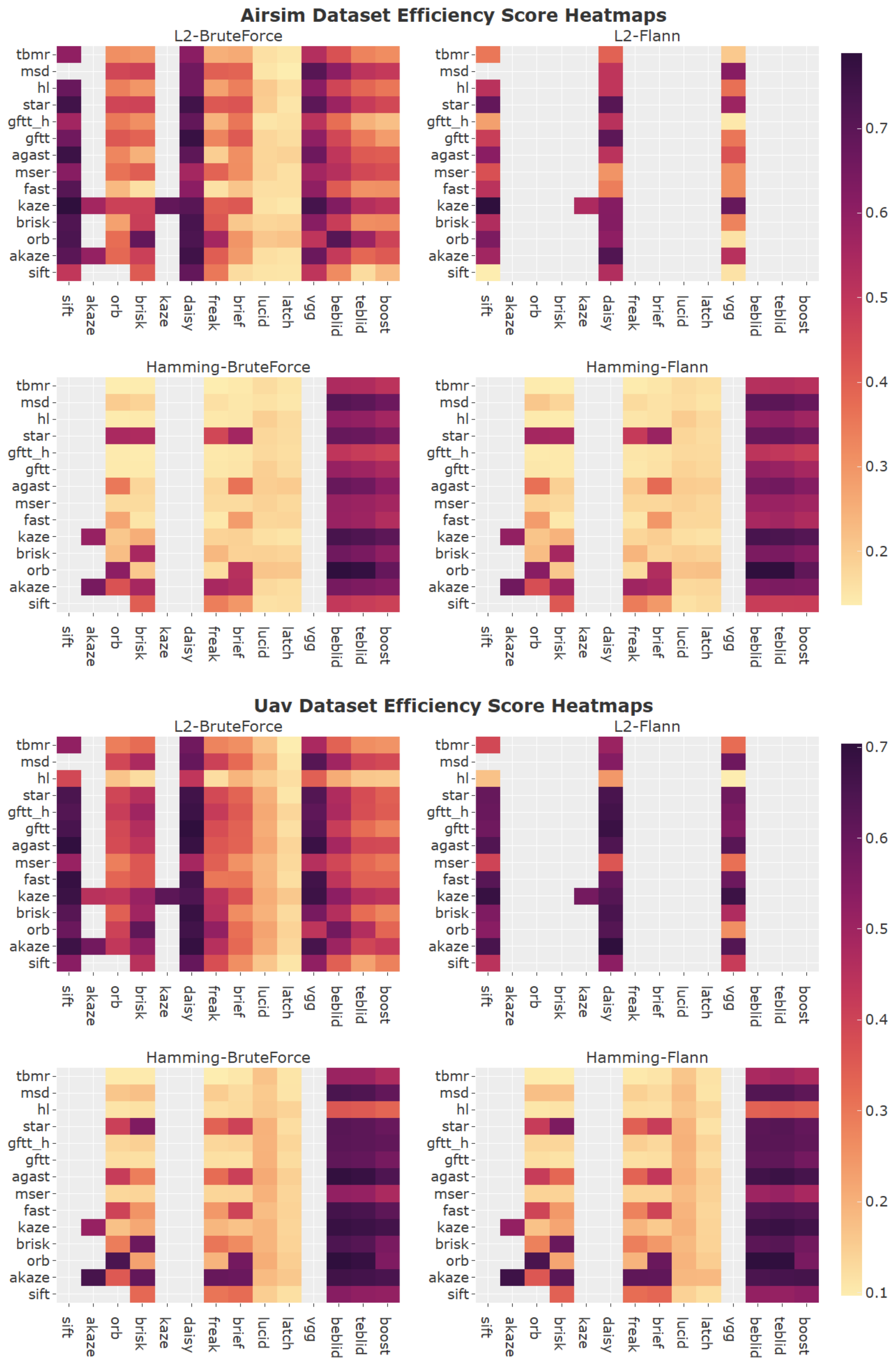
Ammar Abbas ELMAS is researcher at Çukurova University, specializing in UAV-based 3D Reconstruction, Computer Vision, and Image Processing. He graduated with a B.Sc. in Computer Engineering from KTO Karatay University in 2014 and M.Sc. from Çukurova University in 2019. Since 2016, he has been working as a research assistant in the Computer Engineering Department and began his PhD in 2019. He has taught courses labs in the Digital Design, Microprocessors, and Data Structures. His expertise includes 3D mesh compression, feature extraction and matching, and digital image analysis.

APPENDICES

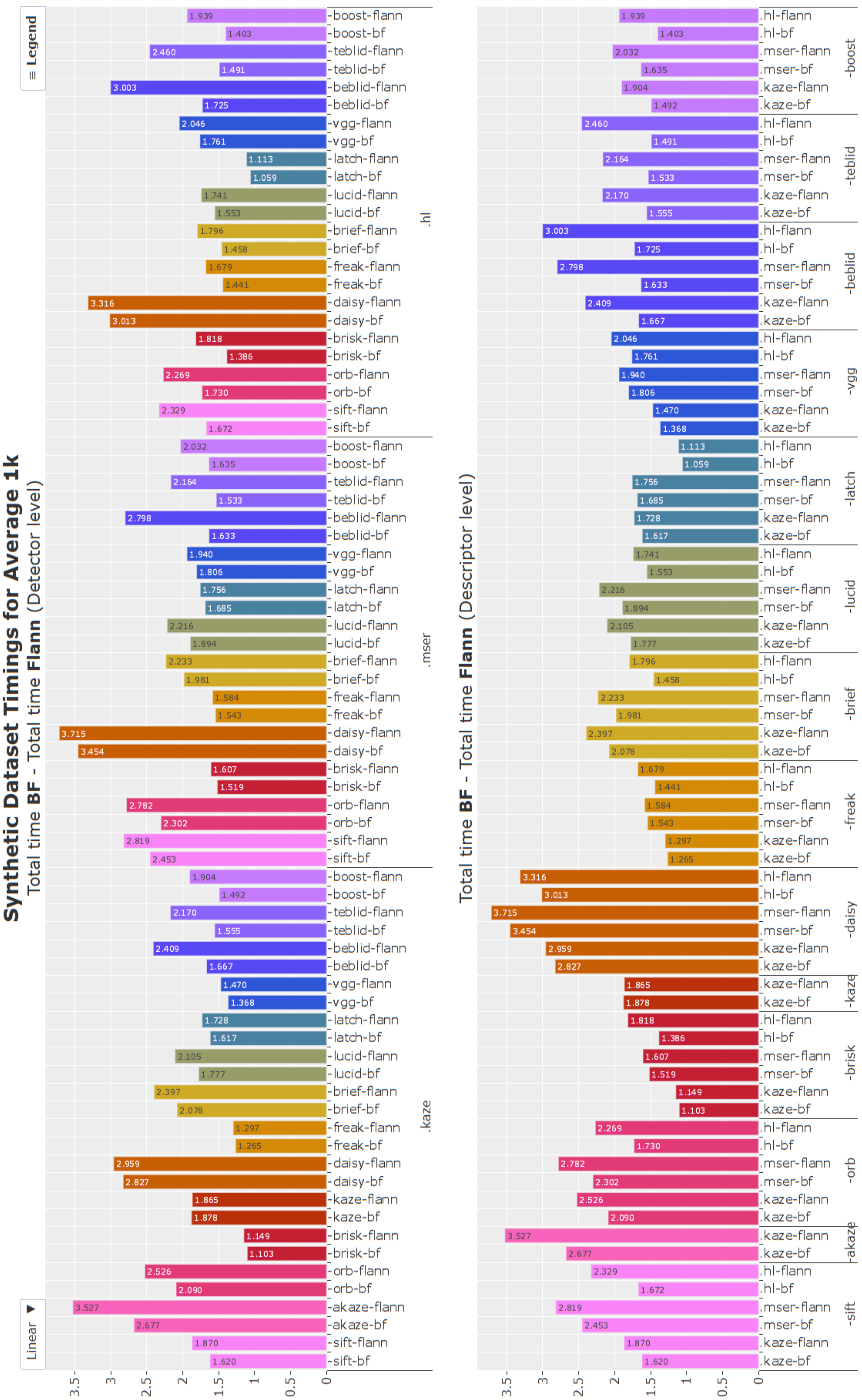
Appendix A: Efficiency Heatmaps

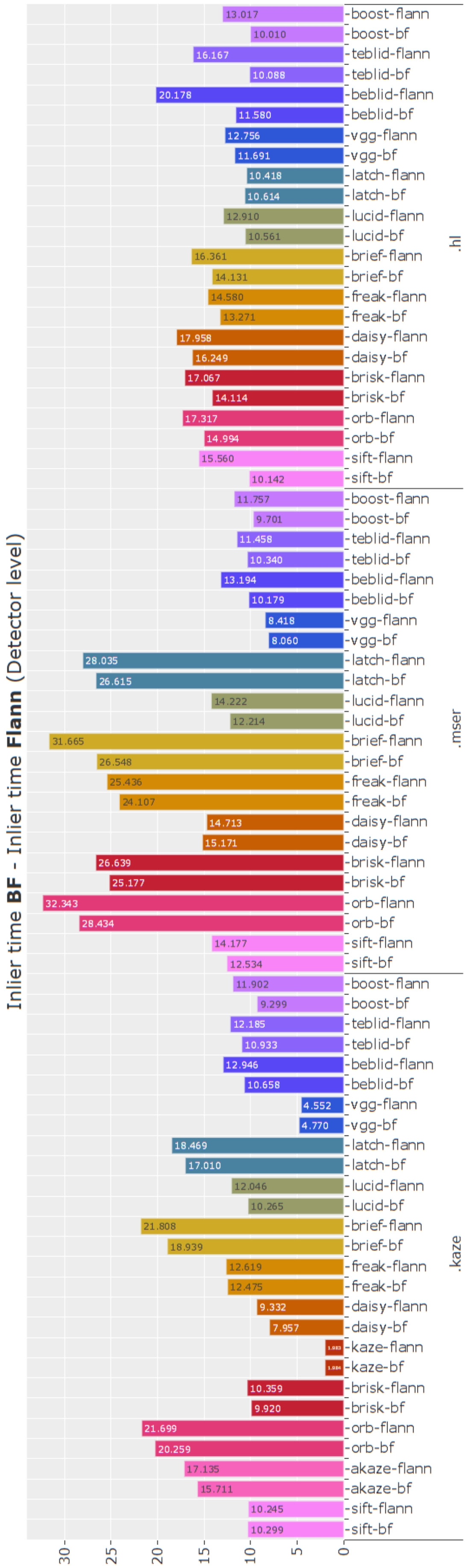


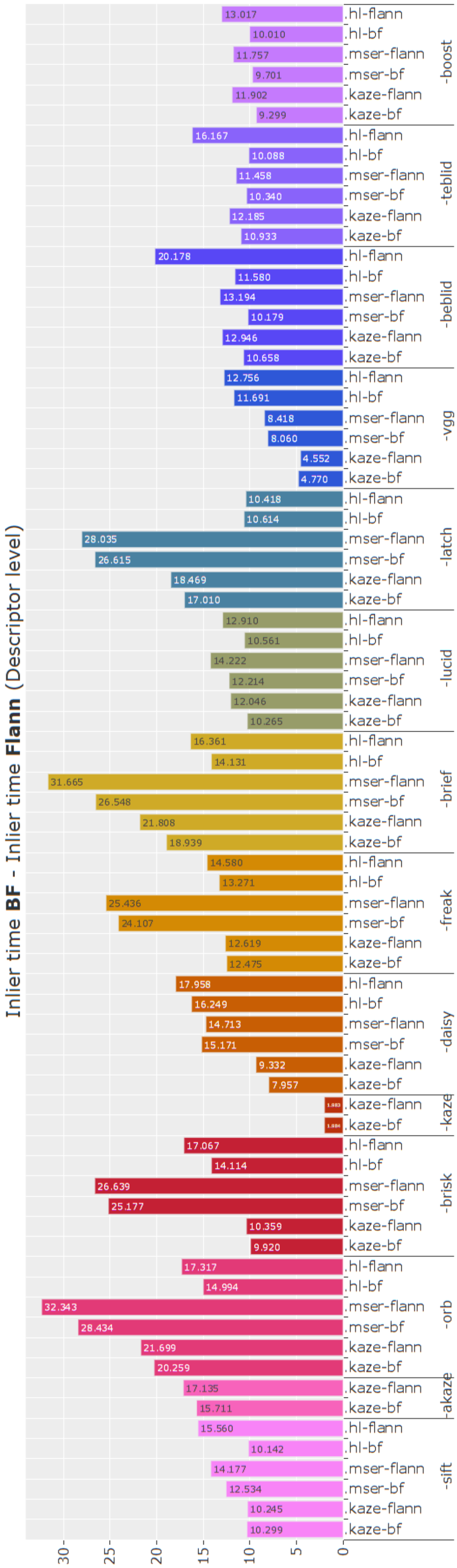



Airsim Dataset Efficiency Score Heatmaps
L2-BruteForce
L2-Flann
Hamming-BruteForce
Hamming-Flann
tbmr
msd
hl
star
gftt_h
gftt
agast
mser
fast
kaze
brisk
orb
akaze
sift
freak
brief
lucid
latch
vgg
daisy
beblid
teblid
boost
0.7
0.6
0.5
0.4
0.3
0.2
Uav Dataset Efficiency Score Heatmaps
0.1

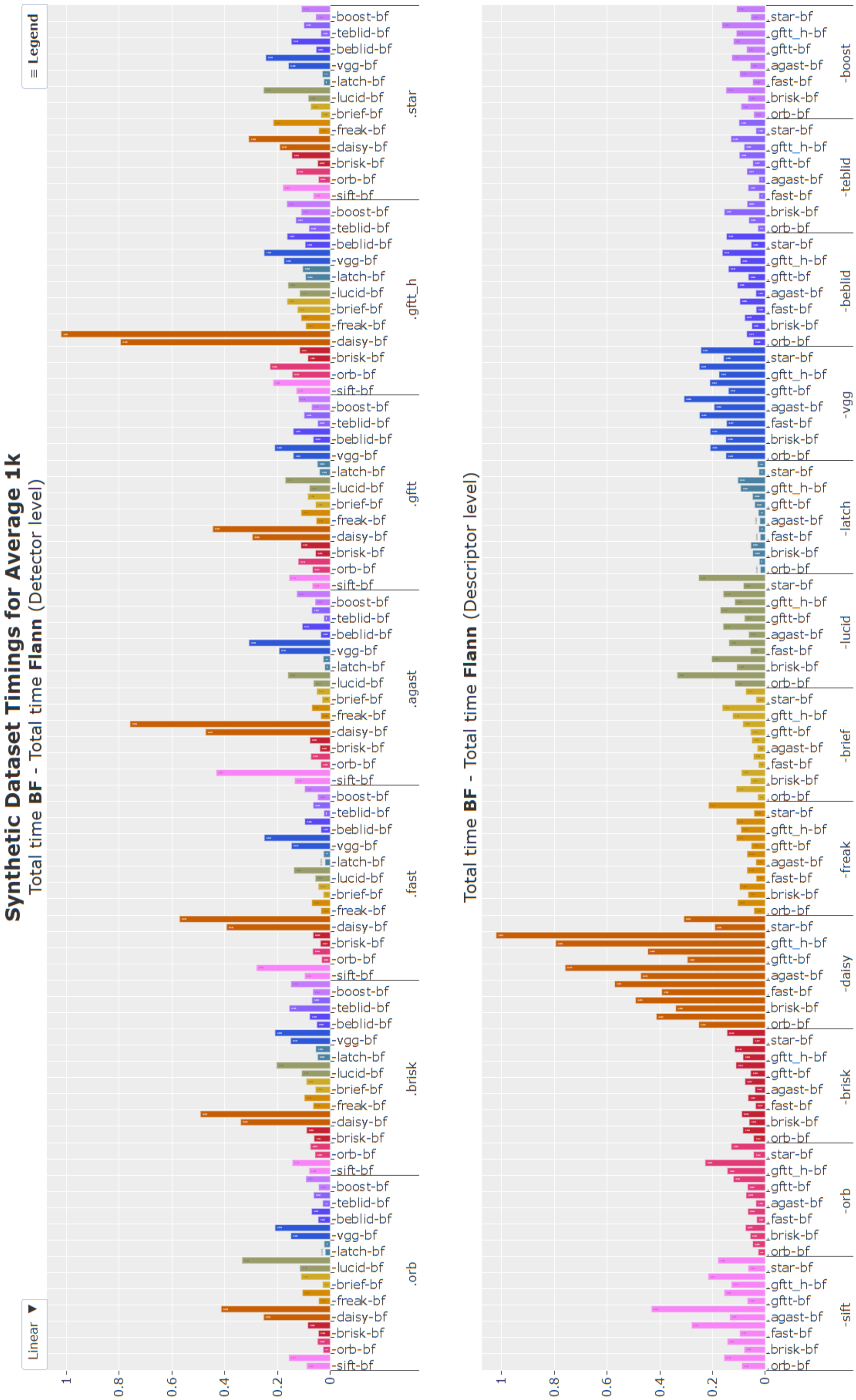
Appendix B.1: Syntethic Dataset Timings of Kaze, Mser, and HL

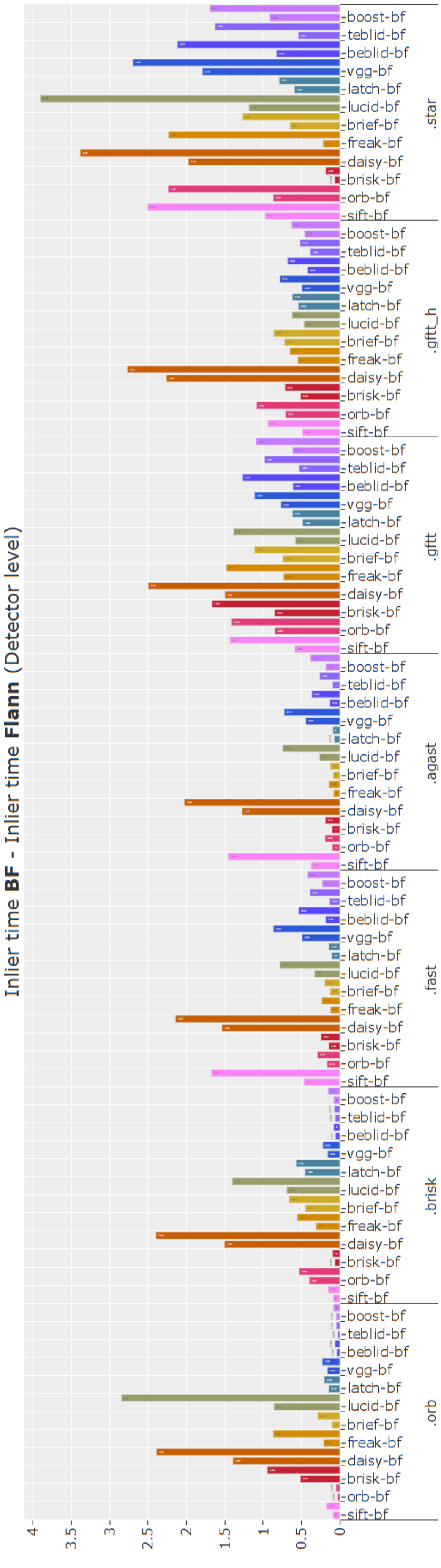


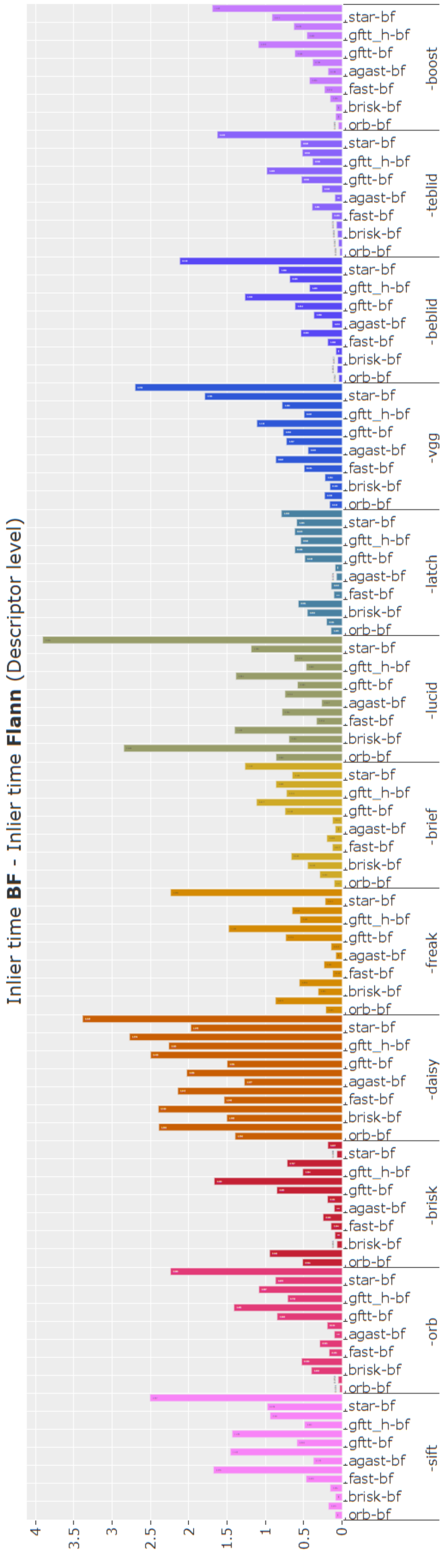

Inlier time BF - Inlier time Flann (Detector level)
.hl
boost-flann 13.017
boost-bf 10.010
teblid-flann 16.167
teblid-bf 10.088
beblid-flann 20.178
beblid-bf 11.580
vgg-flann 12.756
vgg-bf 11.691
latch-flann 10.418
latch-bf 10.614
lucid-flann 12.910
lucid-bf 10.561
brief-flann 16.361
brief-bf 14.131
freak-flann 14.580
freak-bf 13.271
daisy-flann 17.958
daisy-bf 16.249
brisk-flann 17.067
brisk-bf 14.114
orb-flann 17.317
orb-bf 14.994
sift-flann 15.560
sift-bf 10.142
.mser
boost-flann 11.757
boost-bf 9.701
teblid-flann 11.458
teblid-bf 10.340
beblid-flann 13.194
beblid-bf 10.179
vgg-flann 8.418
vgg-bf 8.060
latch-flann 28.035
latch-bf 26.615
lucid-flann 14.222
lucid-bf 12.214
brief-flann 31.665
brief-bf 26.548
freak-flann 25.436
freak-bf 24.107
daisy-flann 14.713
daisy-bf 15.171
brisk-flann 26.639
brisk-bf 25.177
orb-flann 32.343
orb-bf 28.434
sift-flann 14.177
sift-bf 12.534
.kaze
boost-flann 11.902
boost-bf 9.299
teblid-flann 12.185
teblid-bf 10.933
beblid-flann 12.946
beblid-bf 10.658
vgg-flann 4.552
vgg-bf 4.770
latch-flann 18.469
latch-bf 17.010
lucid-flann 12.046
lucid-bf 10.265
brief-flann 21.808
brief-bf 18.939
freak-flann 12.619
freak-bf 12.475
daisy-flann 9.332
daisy-bf 7.957
kaze-flann
kaze-bf
brisk-flann 10.359
brisk-bf 9.920
orb-flann 21.699
orb-bf 20.259
akaze-flann 17.135
akaze-bf 15.711
sift-flann 10.245
sift-bf 10.299
30 25 20 15 10 5 0


Inlier time BF - Inlier time Flann (Descriptor level)
-boost
hl-flann 13.017
hl-bf 10.010
mser-flann 11.757
mser-bf 9.701
kaze-flann 11.902
kaze-bf 9.299
-teblid
hl-flann 16.167
hl-bf 10.088
mser-flann 11.458
mser-bf 10.340
kaze-flann 12.185
kaze-bf 10.933
-beblid
hl-flann 20.178
hl-bf 11.580
mser-flann 13.194
mser-bf 10.179
kaze-flann 12.946
kaze-bf 10.658
-vgg
hl-flann 12.756
hl-bf 11.691
mser-flann 8.418
mser-bf 8.060
kaze-flann 4.552
kaze-bf 4.770
-latch
hl-flann 10.418
hl-bf 10.614
mser-flann 28.035
mser-bf 26.615
kaze-flann 18.469
kaze-bf 17.010
-lucid
hl-flann 12.910
hl-bf 10.561
mser-flann 14.222
mser-bf 12.214
kaze-flann 12.046
kaze-bf 10.265
-brief
hl-flann 16.361
hl-bf 14.131
mser-flann 31.665
mser-bf 26.548
kaze-flann 21.808
kaze-bf 18.939
-freak
hl-flann 14.580
hl-bf 13.271
mser-flann 25.436
mser-bf 24.107
kaze-flann 12.619
kaze-bf 12.475
-daisy
hl-flann 17.958
hl-bf 16.249
mser-flann 14.713
mser-bf 15.171
kaze-flann 9.332
kaze-bf 7.957
-kaze
kaze-flann
kaze-bf
-brisk
hl-flann 17.067
hl-bf 14.114
mser-flann 26.639
mser-bf 25.177
kaze-flann 10.359
kaze-bf 9.920
-orb
hl-flann 17.317
hl-bf 14.994
mser-flann 32.343
mser-bf 28.434
kaze-flann 21.699
kaze-bf 20.259
-akaze
kaze-flann 17.135
kaze-bf 15.711
-sift
hl-flann 15.560
hl-bf 10.142
mser-flann 14.177
mser-bf 12.534
kaze-flann 10.245
kaze-bf 10.299
30 25 20 15 10 5 0

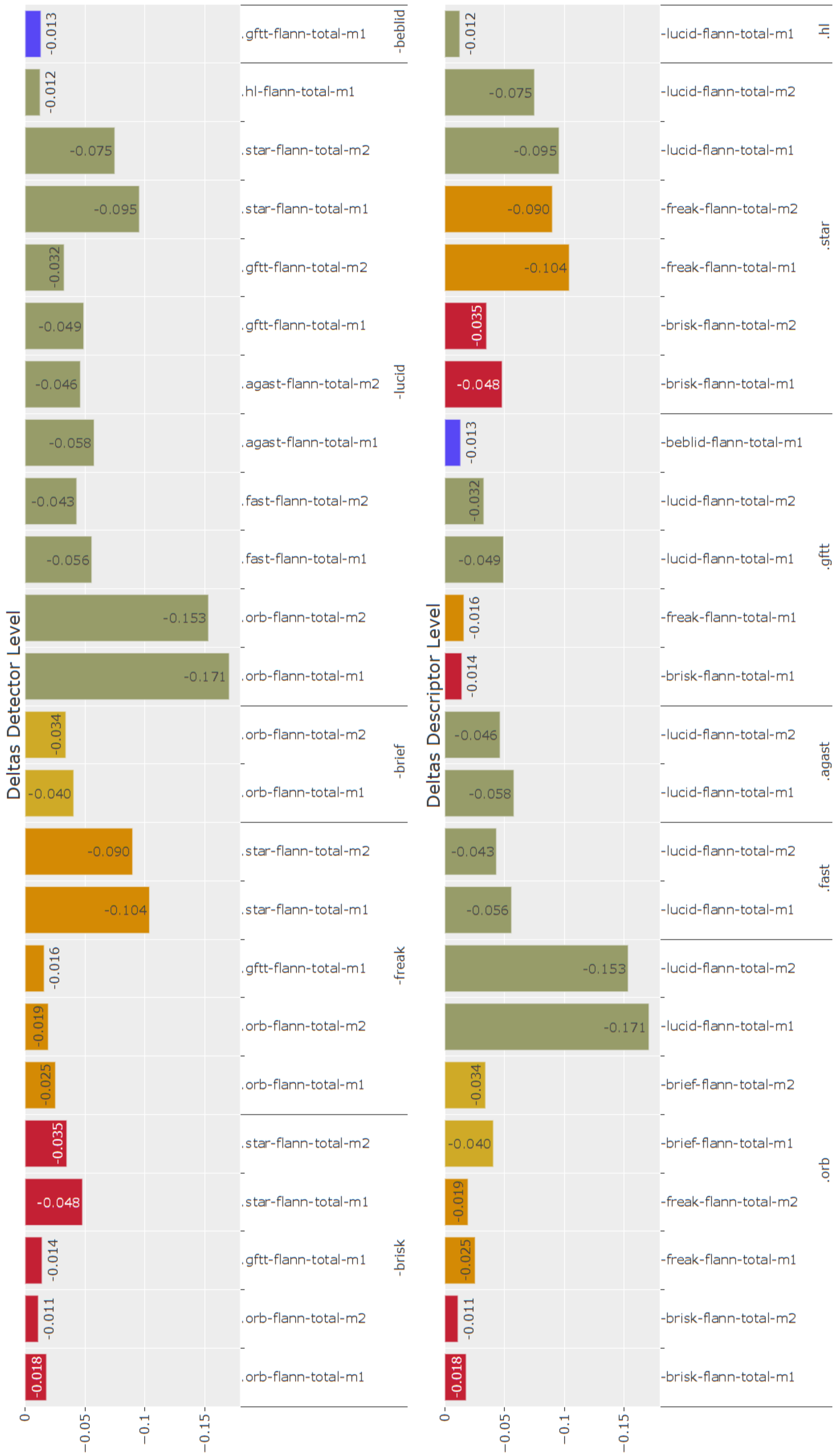
Appendix B.2: Synthetic Dataset Timings of Orb, Brisk, Fast, Agast, Gftt, and Star

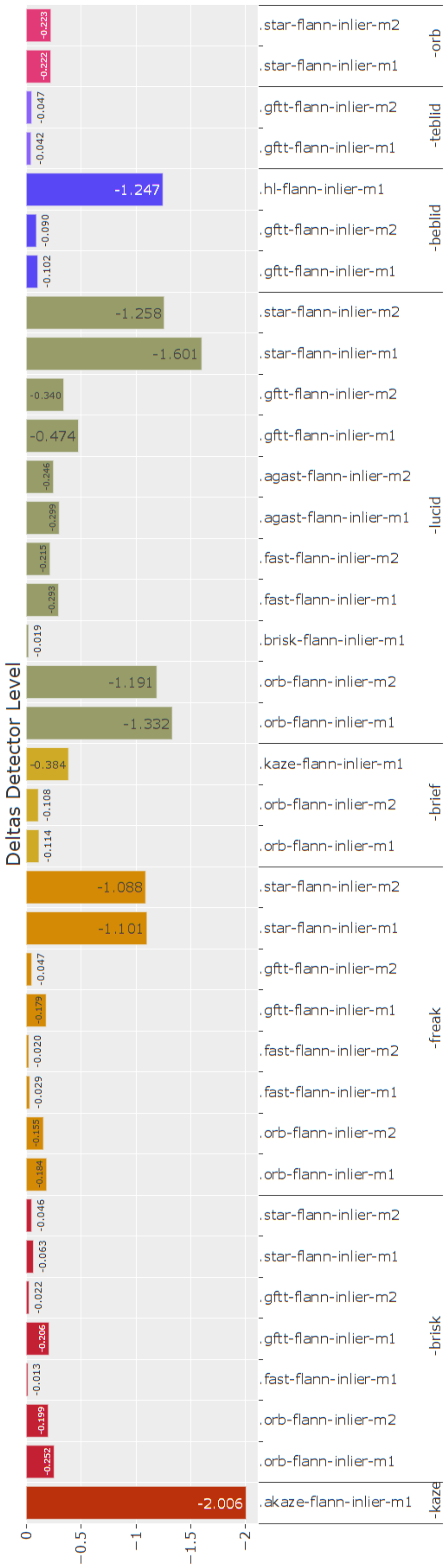


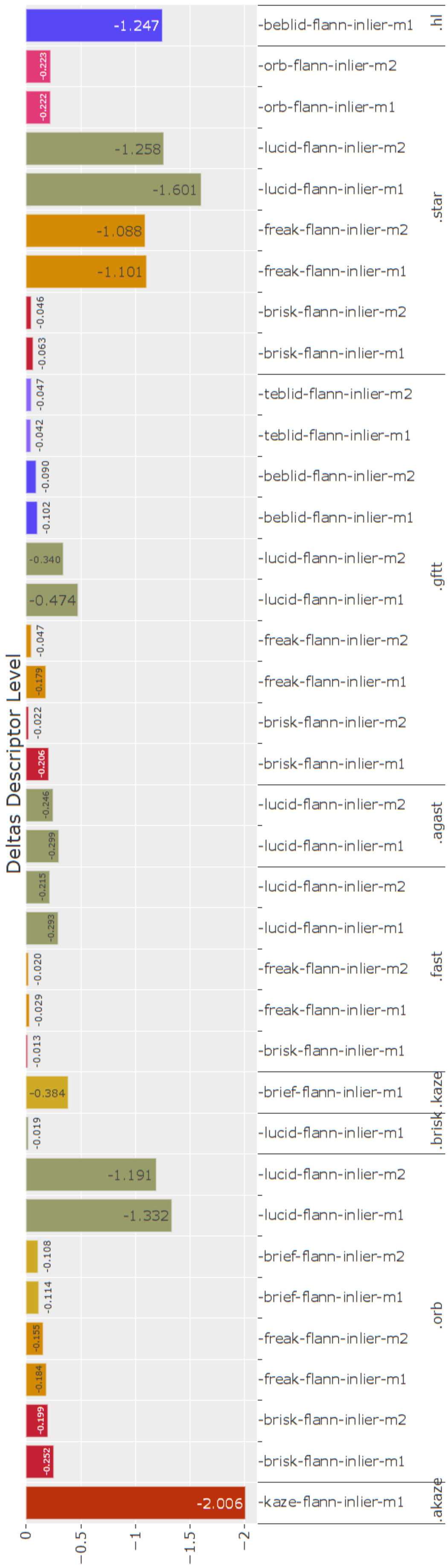

Inlier time BF - Inlier time Flann (Detector level)
star
gftt_h
gftt
agast
fast
brisk
orb


Inlier time BF - Inlier time Flann (Descriptor level)
boost
teblid
beblid
vgg
latch
lucid
brief
freak
daisy
brisk
orb
sift

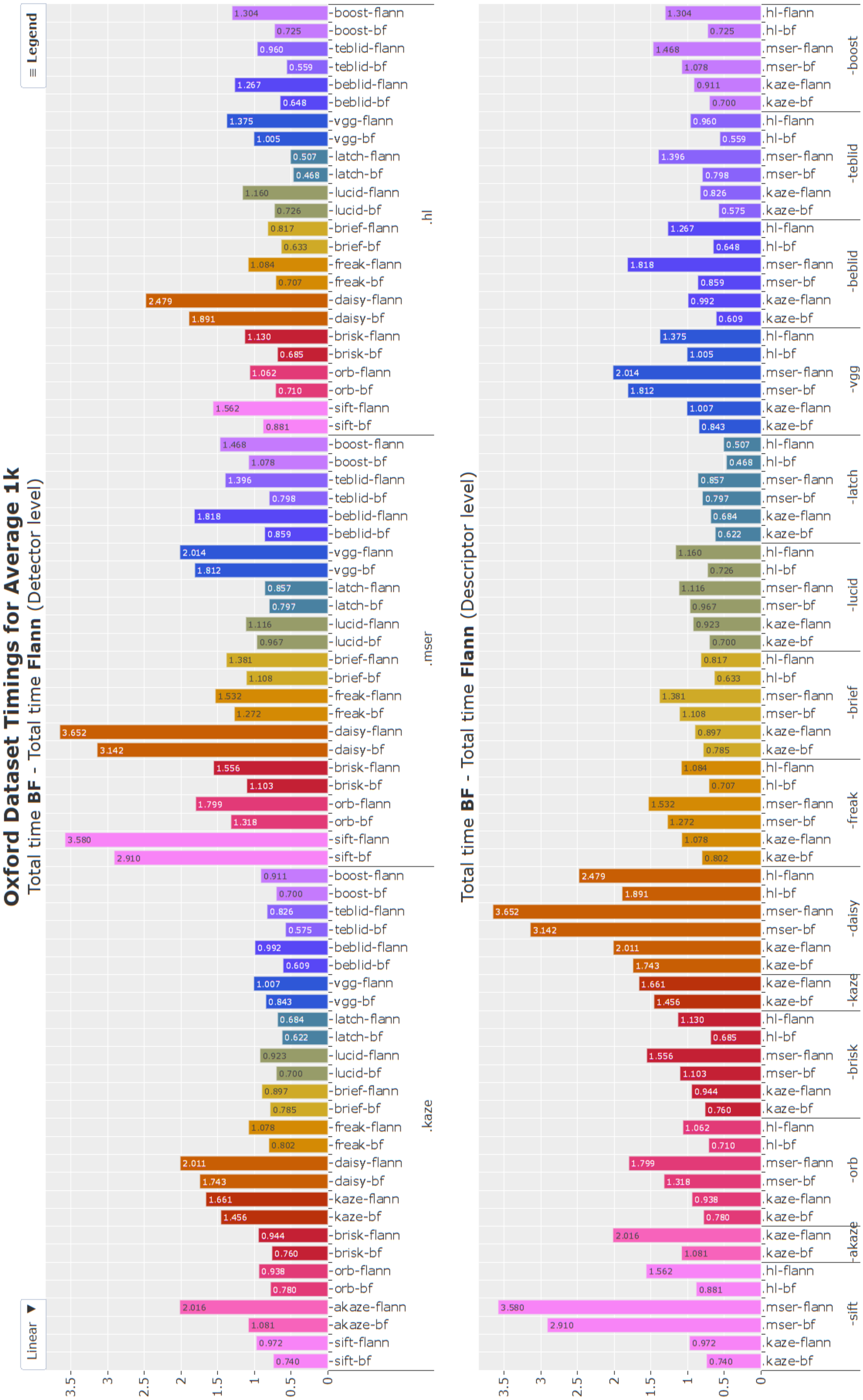
Appendix B.3: Synthetic Dataset Faster Mobile Timings (Mobile - Server)

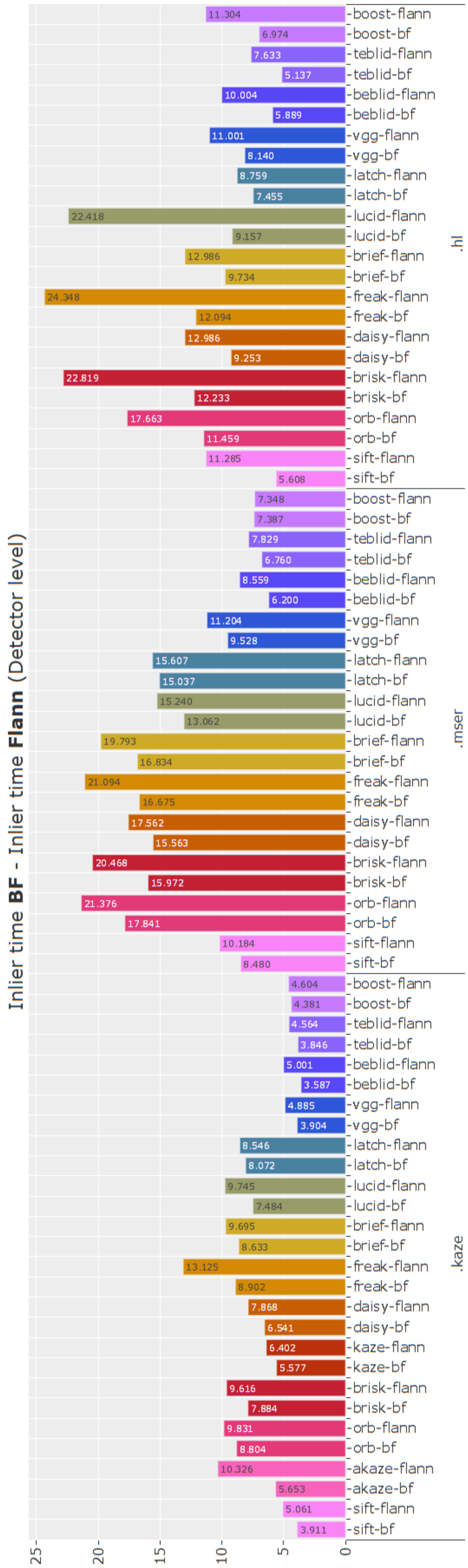


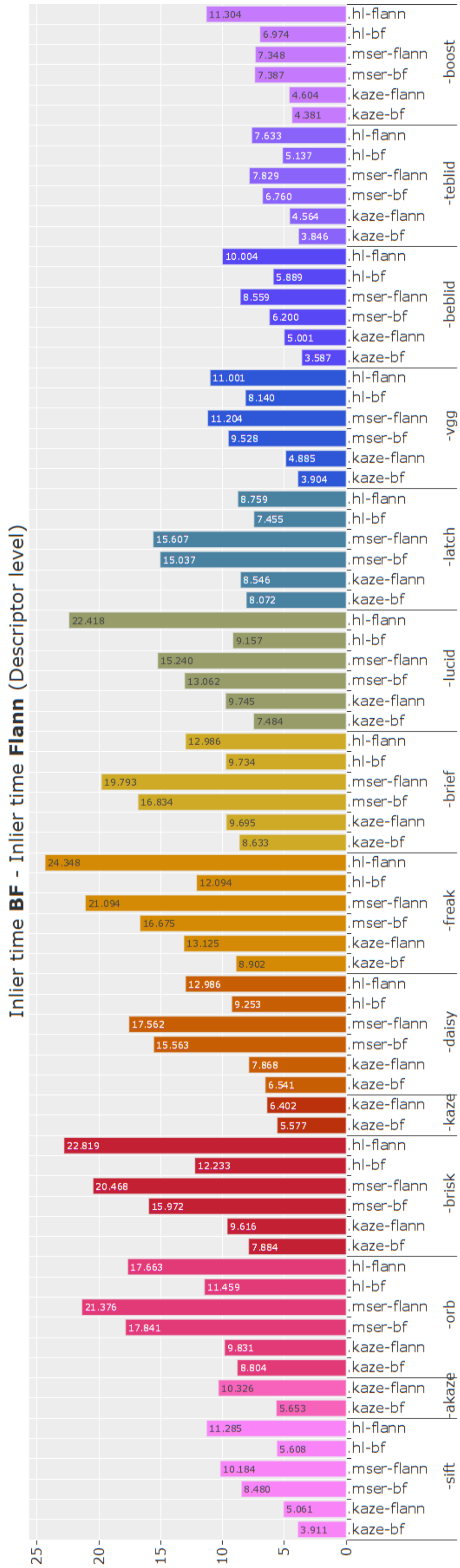

Deltas Detector Level


Deltas Descriptor Level

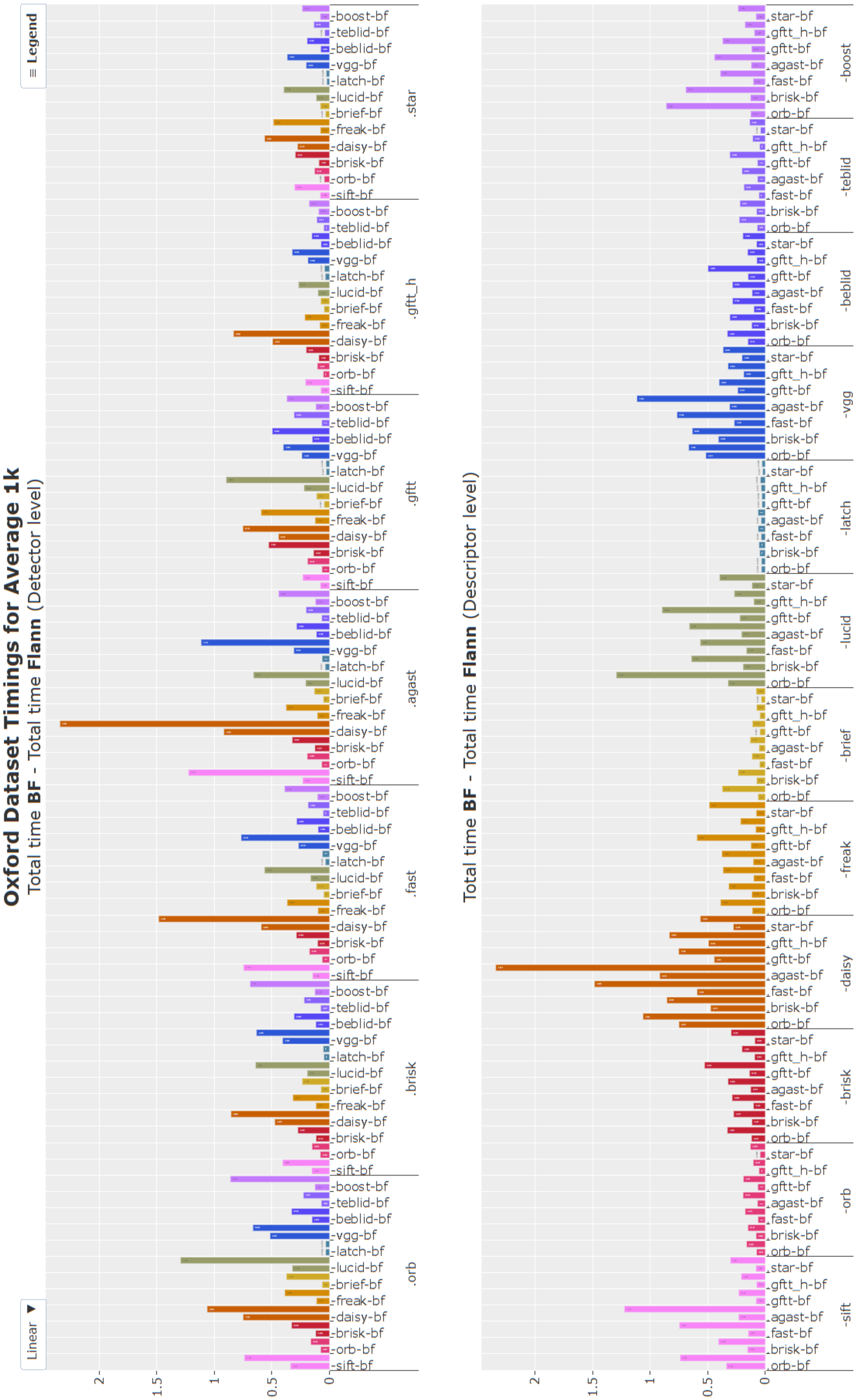
Appendix C.1: Oxford Dataset Timings of Kaze, Mser, and HL

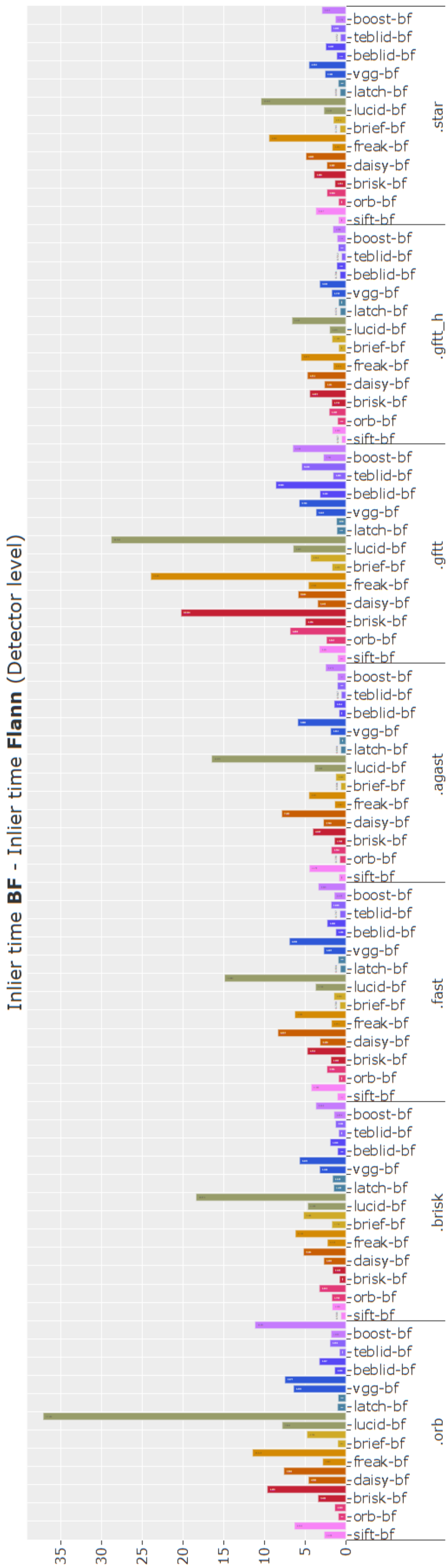


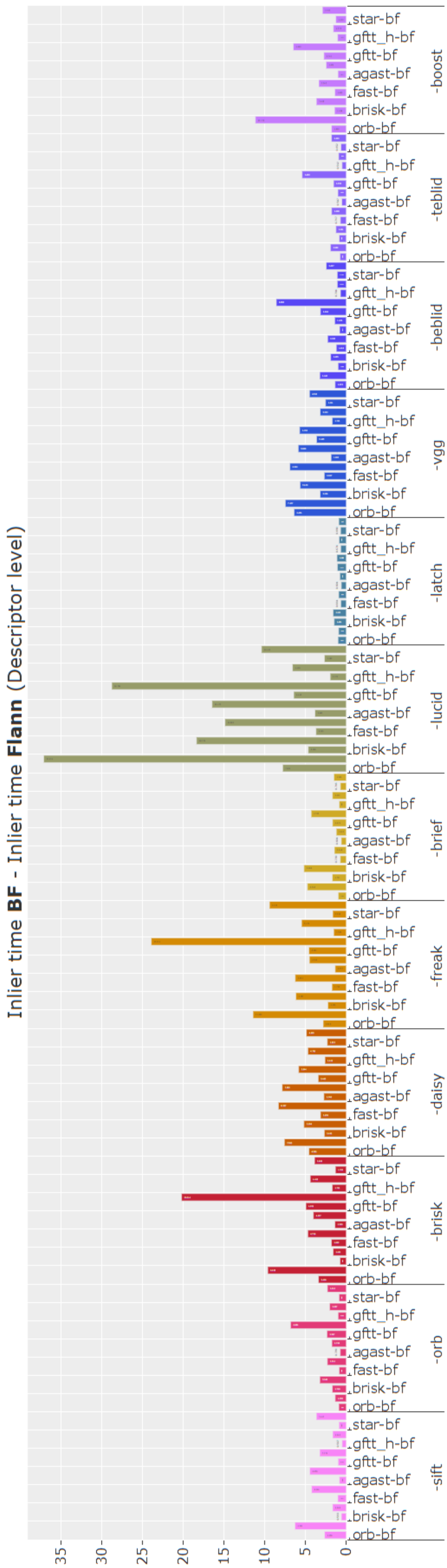

Inlier time BF - Inlier time Flann (Detector level)
.hl
.mser
.kaze
25
20
15
10
5
0


Inlier time BF - Inlier time Flann (Descriptor level)
-boost
-teblid
-beblid
-vgg
-latch
-lucid
-brief
-freak
-daisy
-kaze
-brisk
-orb
-akaze
-sift
25
20
15
10
5
0

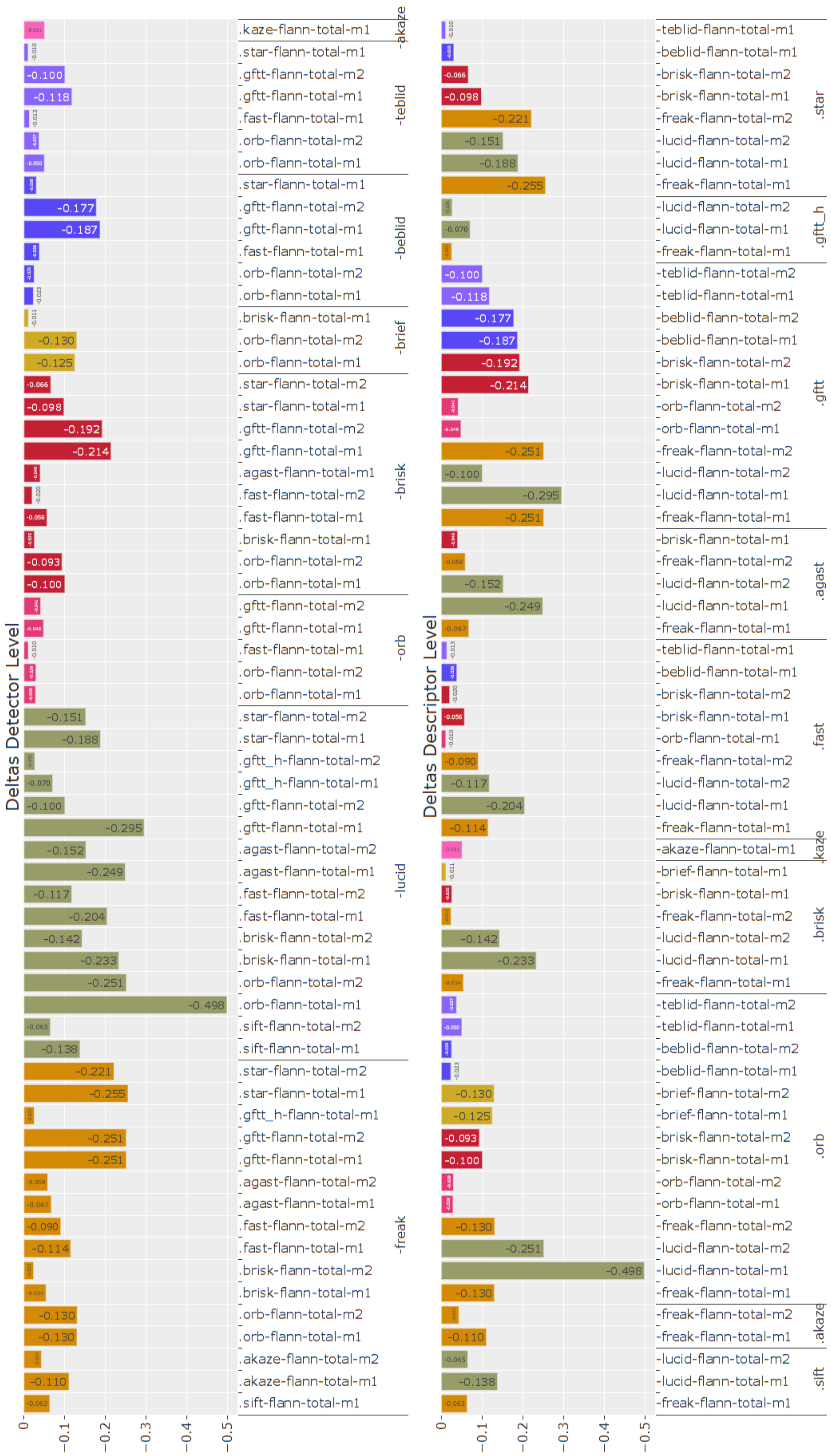
Appendix C.2: Oxford Dataset Timings of Orb, Brisk, Fast, Agast, Gftt, and Star

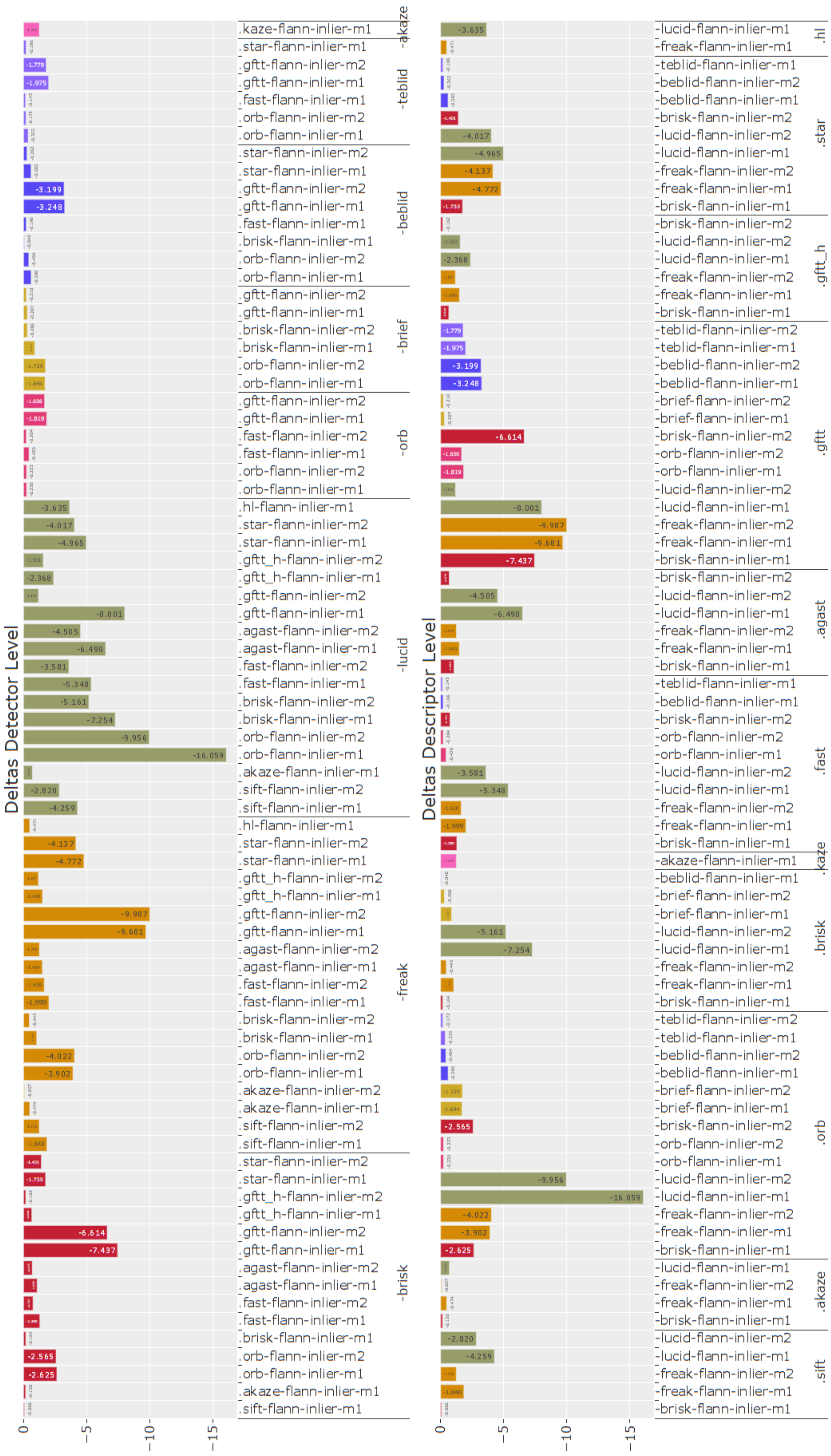



Inlier time BF - Inlier time Flann (Detector level)
star
gftt_h
gftt
agast
fast
brisk
orb
boost-bf
teblid-bf
beblid-bf
vgg-bf
latch-bf
lucid-bf
brief-bf
freak-bf
daisy-bf
brisk-bf
orb-bf
sift-bf
35
30
25
20
15
10
5
0


Inlier time BF - Inlier time Flann (Descriptor level)
boost
teblid
beblid
vgg
latch
lucid
brief
freak
daisy
brisk
orb
sift
star-bf
gftt_h-bf
gftt-bf
agast-bf
fast-bf
brisk-bf
orb-bf
35
30
25
20
15
10
5
0

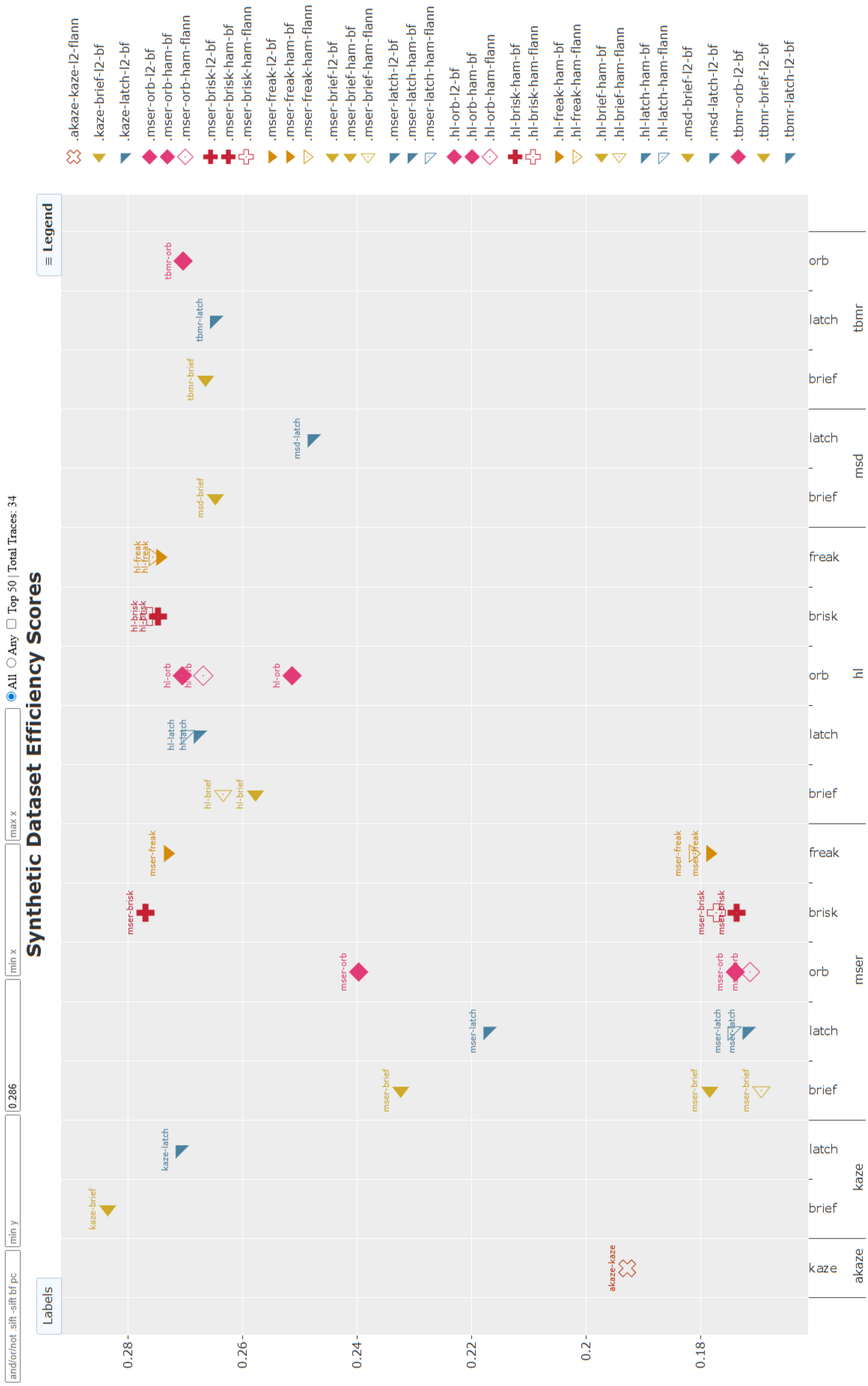
Appendix C.3: Oxford Dataset Faster Mobile Timings (Mobile - Server)

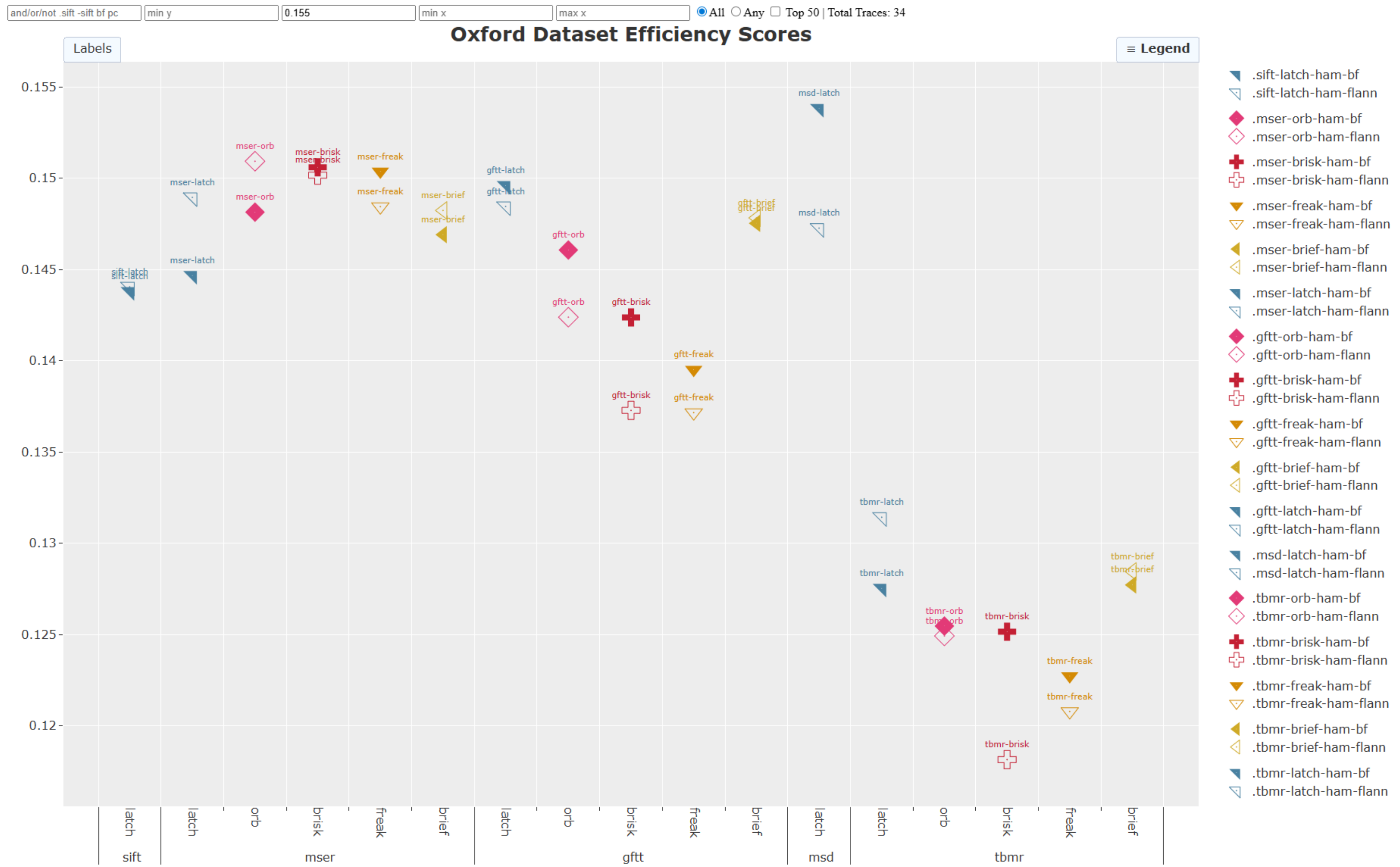


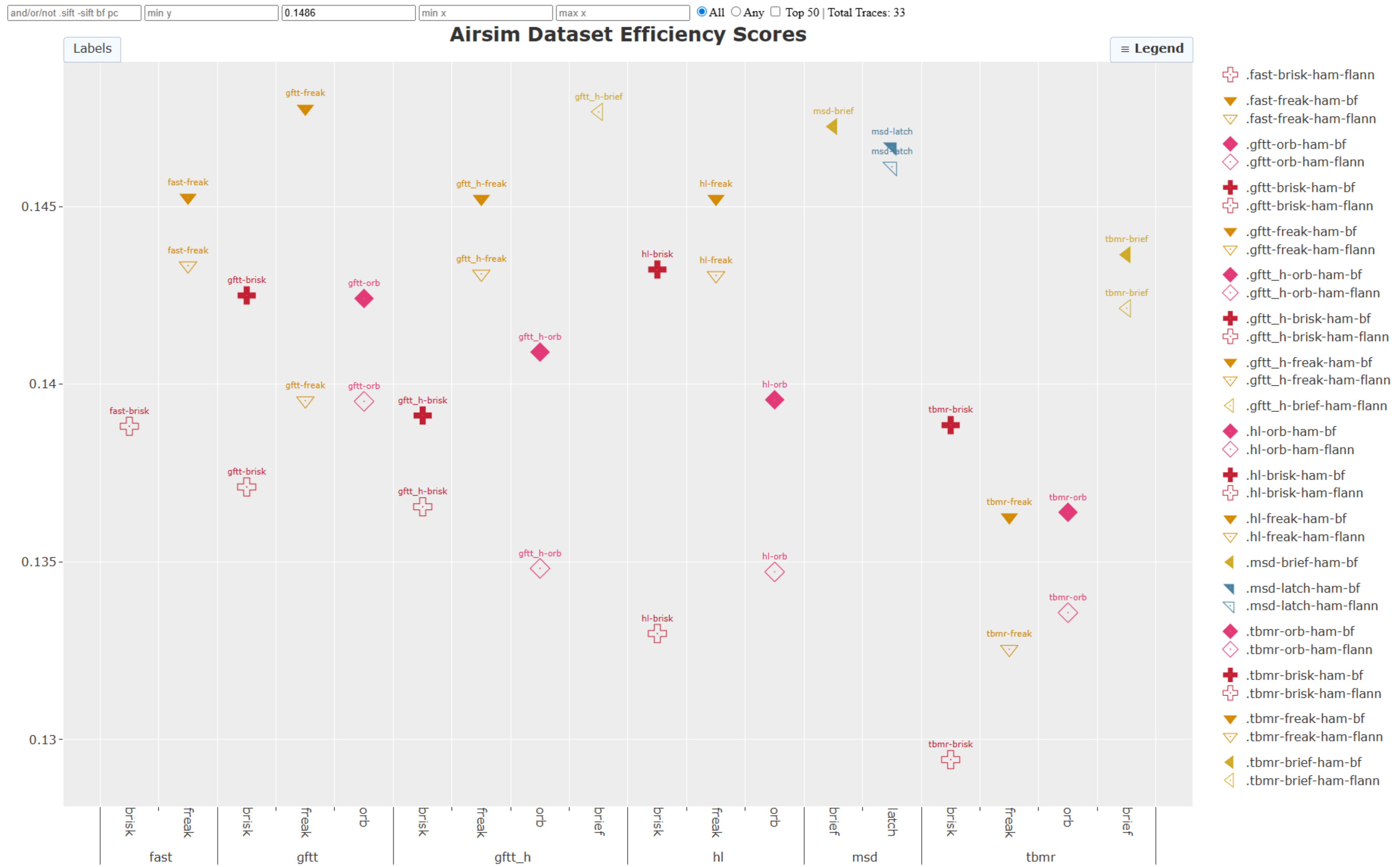


Deltas Detector Level
Deltas Descriptor Level

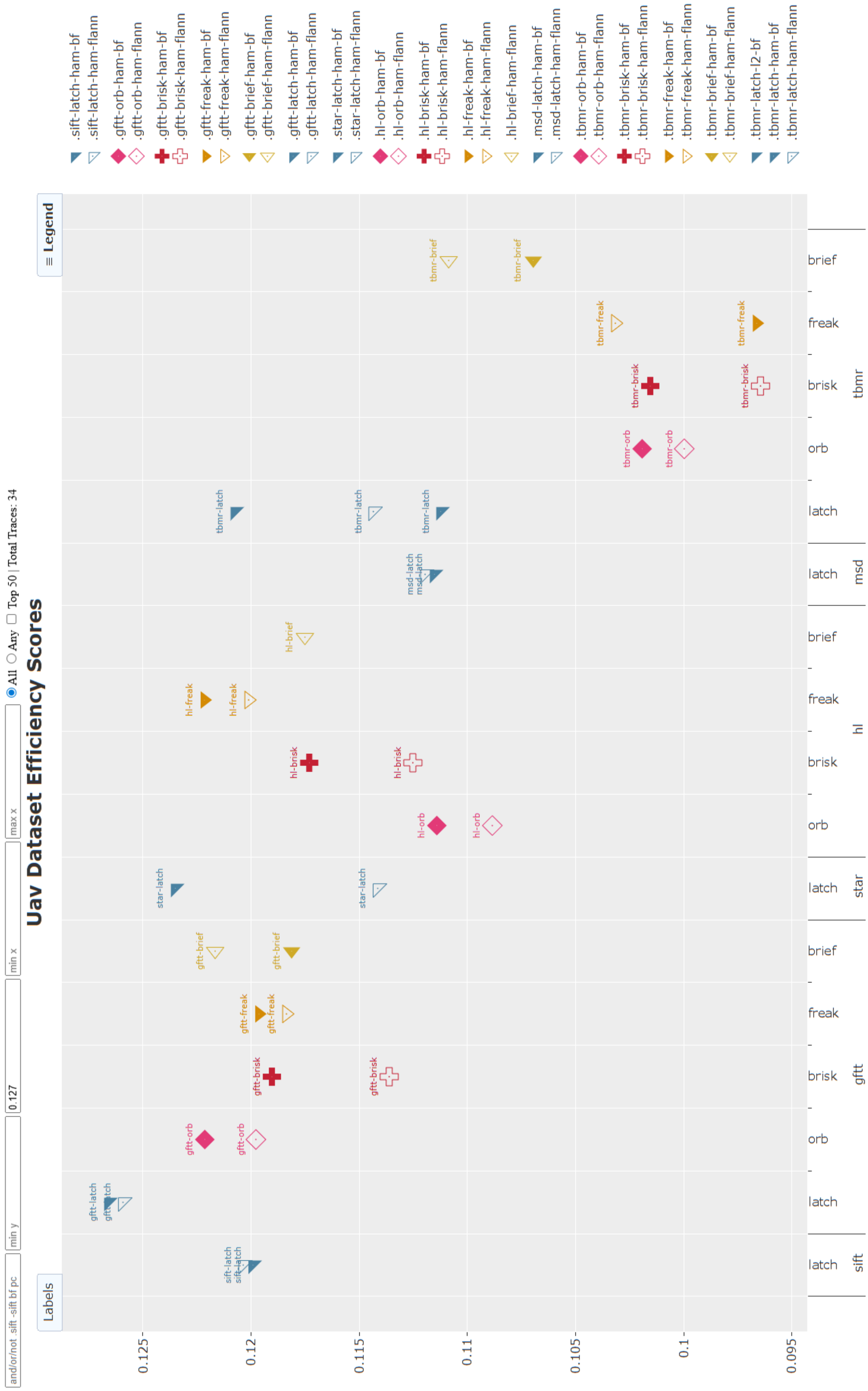
Appendix D.1: Worst Efficiency Scores



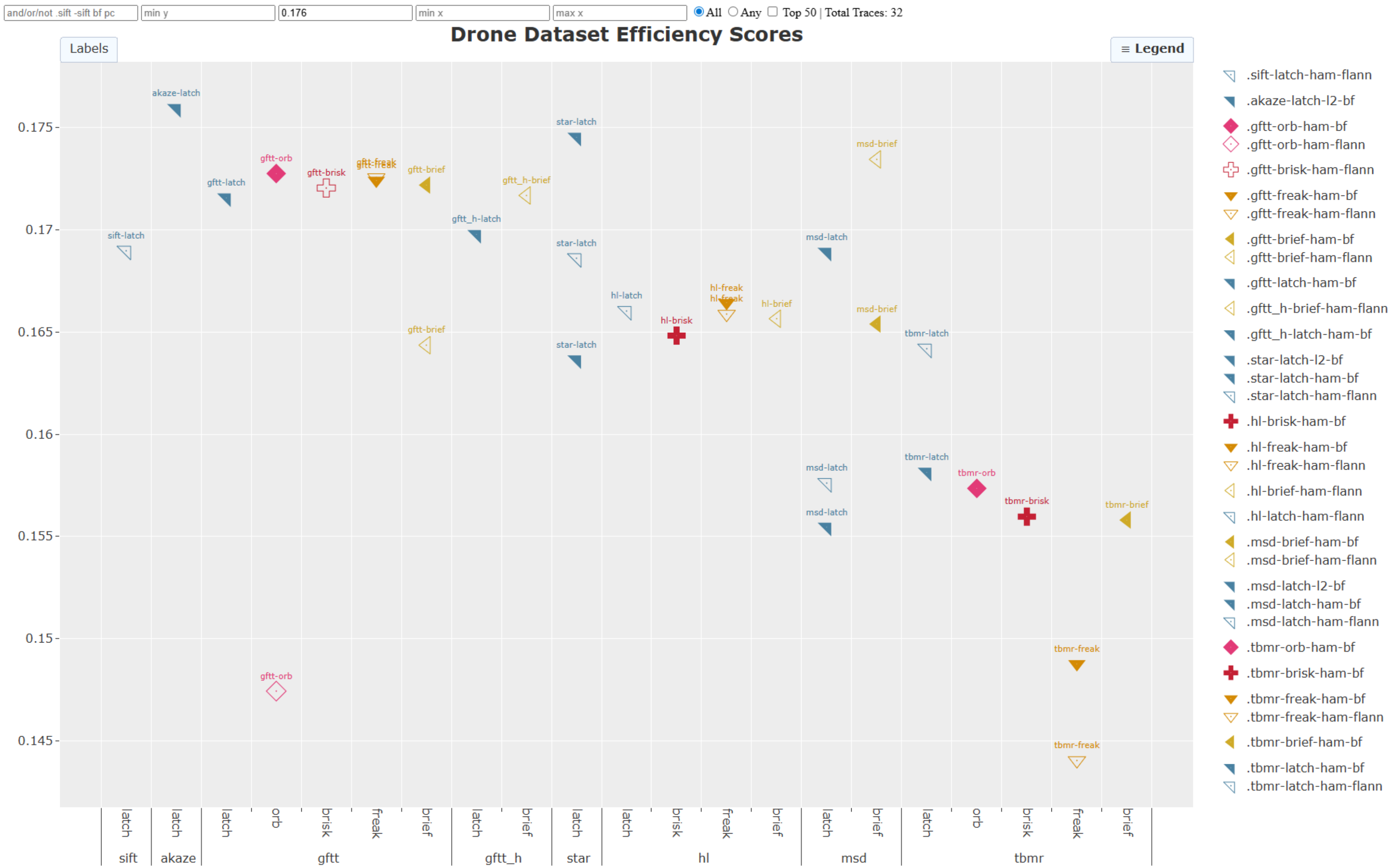

and/or/not .sift -sift bf pc
min y
0.155
min x
max x
All
Any
Top 50 | Total Traces: 34
Oxford Dataset Efficiency Scores
Labels
≡ Legend
0.155
0.15
0.145
0.14
0.135
0.13
0.125
0.12
latch
sift
latch
orb
brisk
freak
brief
mser
latch
orb
brisk
freak
brief
gftt
latch
msd
latch
orb
brisk
freak
brief
tbmr
.sift-latch-ham-bf
.sift-latch-ham-flann
.mser-orb-ham-bf
.mser-orb-ham-flann
.mser-brisk-ham-bf
.mser-brisk-ham-flann
.mser-freak-ham-bf
.mser-freak-ham-flann
.mser-brief-ham-bf
.mser-brief-ham-flann
.mser-latch-ham-bf
.mser-latch-ham-flann
.gftt-orb-ham-bf
.gftt-orb-ham-flann
.gftt-brisk-ham-bf
.gftt-brisk-ham-flann
.gftt-freak-ham-bf
.gftt-freak-ham-flann
.gftt-brief-ham-bf
.gftt-brief-ham-flann
.gftt-latch-ham-bf
.gftt-latch-ham-flann
.msd-latch-ham-bf
.msd-latch-ham-flann
.tbmr-orb-ham-bf
.tbmr-orb-ham-flann
.tbmr-brisk-ham-bf
.tbmr-brisk-ham-flann
.tbmr-freak-ham-bf
.tbmr-freak-ham-flann
.tbmr-brief-ham-bf
.tbmr-brief-ham-flann
.tbmr-latch-ham-bf
.tbmr-latch-ham-flann



Airsim Dataset Efficiency Scores
and/or/not .sift -sift bf pc
min y
0.1486
min x
max x
All
Any
Top 50 | Total Traces: 33
Labels
Legend
0.145
0.14
0.135
0.13
brisk
freak
orb
brief
latch
fast
gftt
gftt_h
hl
msd
tbmr
.fast-brisk-ham-flann
.fast-freak-ham-bf
.fast-freak-ham-flann
.gftt-orb-ham-bf
.gftt-orb-ham-flann
.gftt-brisk-ham-bf
.gftt-brisk-ham-flann
.gftt-freak-ham-bf
.gftt-freak-ham-flann
.gftt_h-orb-ham-bf
.gftt_h-orb-ham-flann
.gftt_h-brisk-ham-bf
.gftt_h-brisk-ham-flann
.gftt_h-freak-ham-bf
.gftt_h-freak-ham-flann
.gftt_h-brief-ham-flann
.hl-orb-ham-bf
.hl-orb-ham-flann
.hl-brisk-ham-bf
.hl-brisk-ham-flann
.hl-freak-ham-bf
.hl-freak-ham-flann
.msd-brief-ham-bf
.msd-latch-ham-bf
.msd-latch-ham-flann
.tbmr-orb-ham-bf
.tbmr-orb-ham-flann
.tbmr-brisk-ham-bf
.tbmr-brisk-ham-flann
.tbmr-freak-ham-bf
.tbmr-freak-ham-flann
.tbmr-brief-ham-bf
.tbmr-brief-ham-flann

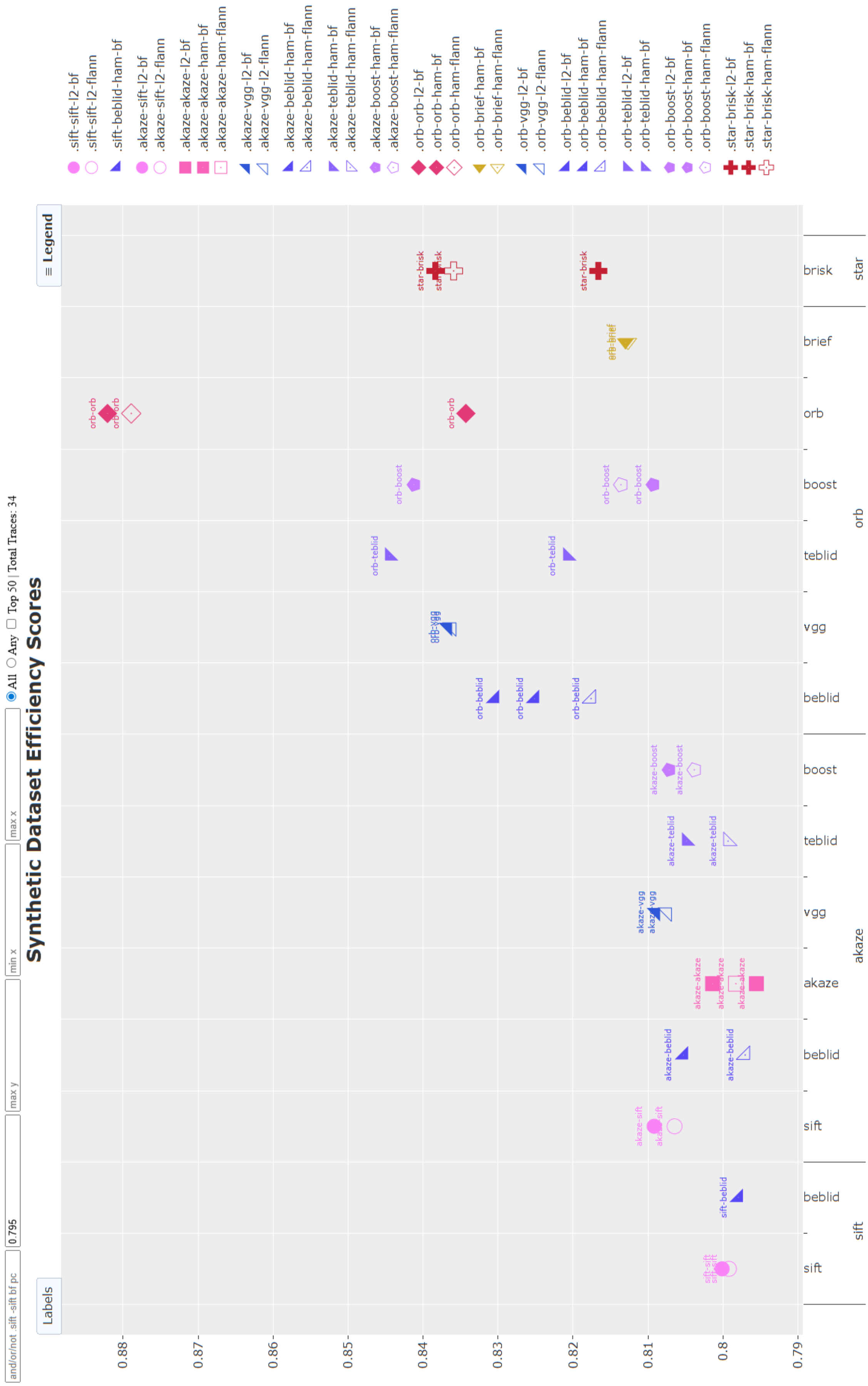


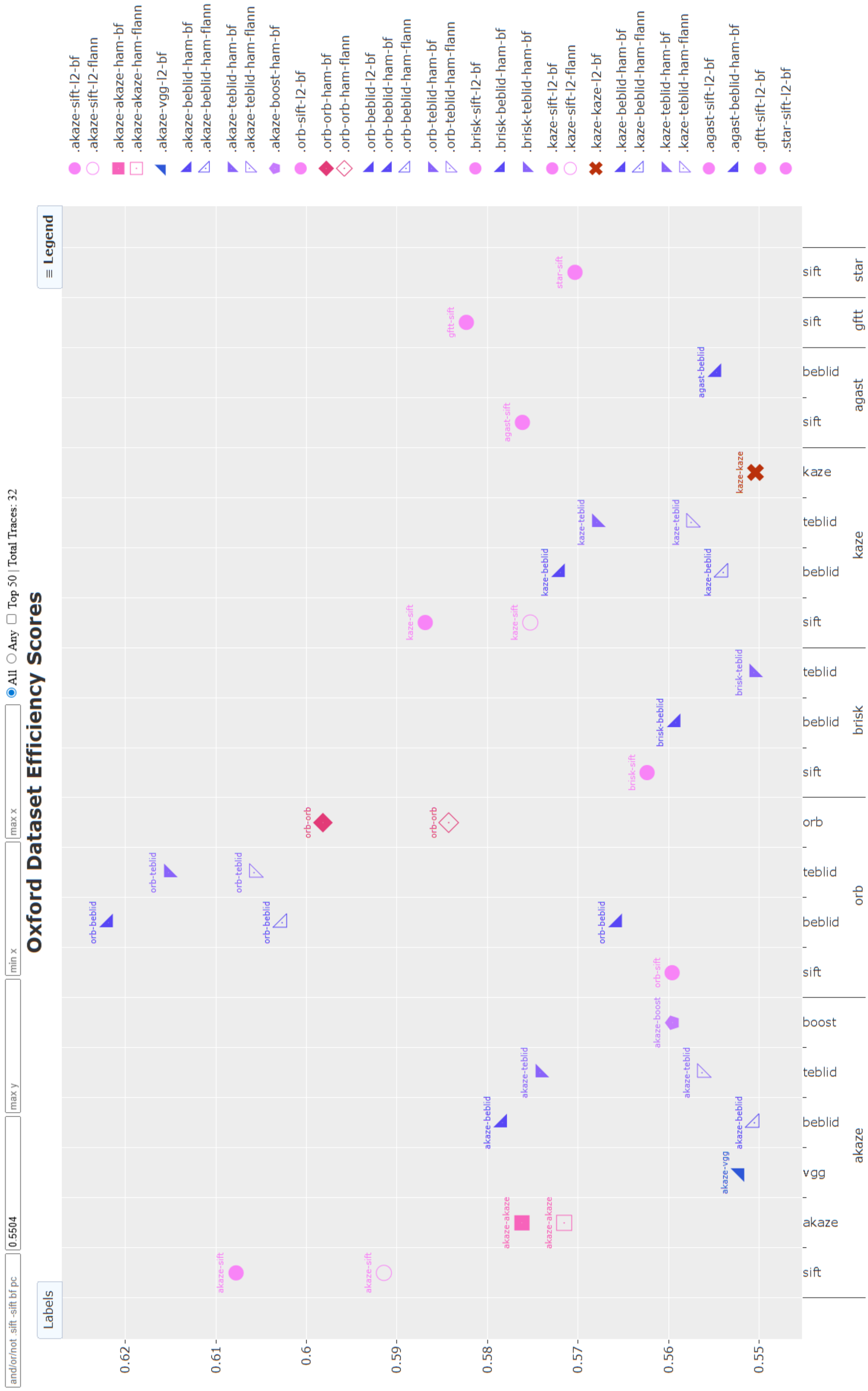
Uav Dataset Efficiency Scores
and/or/not .sift -sift bf pc
min y
0.127
min x
max x
All Any Top 50 | Total Traces: 34
Labels
≡ Legend
.sift-latch-ham-bf
.sift-latch-ham-flann
.gftt-orb-ham-bf
.gftt-orb-ham-flann
.gftt-brisk-ham-bf
.gftt-brisk-ham-flann
.gftt-freak-ham-bf
.gftt-freak-ham-flann
.gftt-brief-ham-bf
.gftt-brief-ham-flann
.gftt-latch-ham-bf
.gftt-latch-ham-flann
.star-latch-ham-bf
.star-latch-ham-flann
.hl-orb-ham-bf
.hl-orb-ham-flann
.hl-brisk-ham-bf
.hl-brisk-ham-flann
.hl-freak-ham-bf
.hl-freak-ham-flann
.hl-brief-ham-flann
.msd-latch-ham-bf
.msd-latch-ham-flann
.tbmr-orb-ham-bf
.tbmr-orb-ham-flann
.tbmr-brisk-ham-bf
.tbmr-brisk-ham-flann
.tbmr-freak-ham-bf
.tbmr-freak-ham-flann
.tbmr-brief-ham-bf
.tbmr-brief-ham-flann
.tbmr-latch-l2-bf
.tbmr-latch-ham-bf
.tbmr-latch-ham-flann
0.125
0.12
0.115
0.11
0.105
0.1
0.095
latch
sift
latch
orb
brisk
freak
brief
gftt
latch
star
orb
brisk
freak
brief
hl
latch
msd
latch
orb
brisk
freak
brief
tbmr

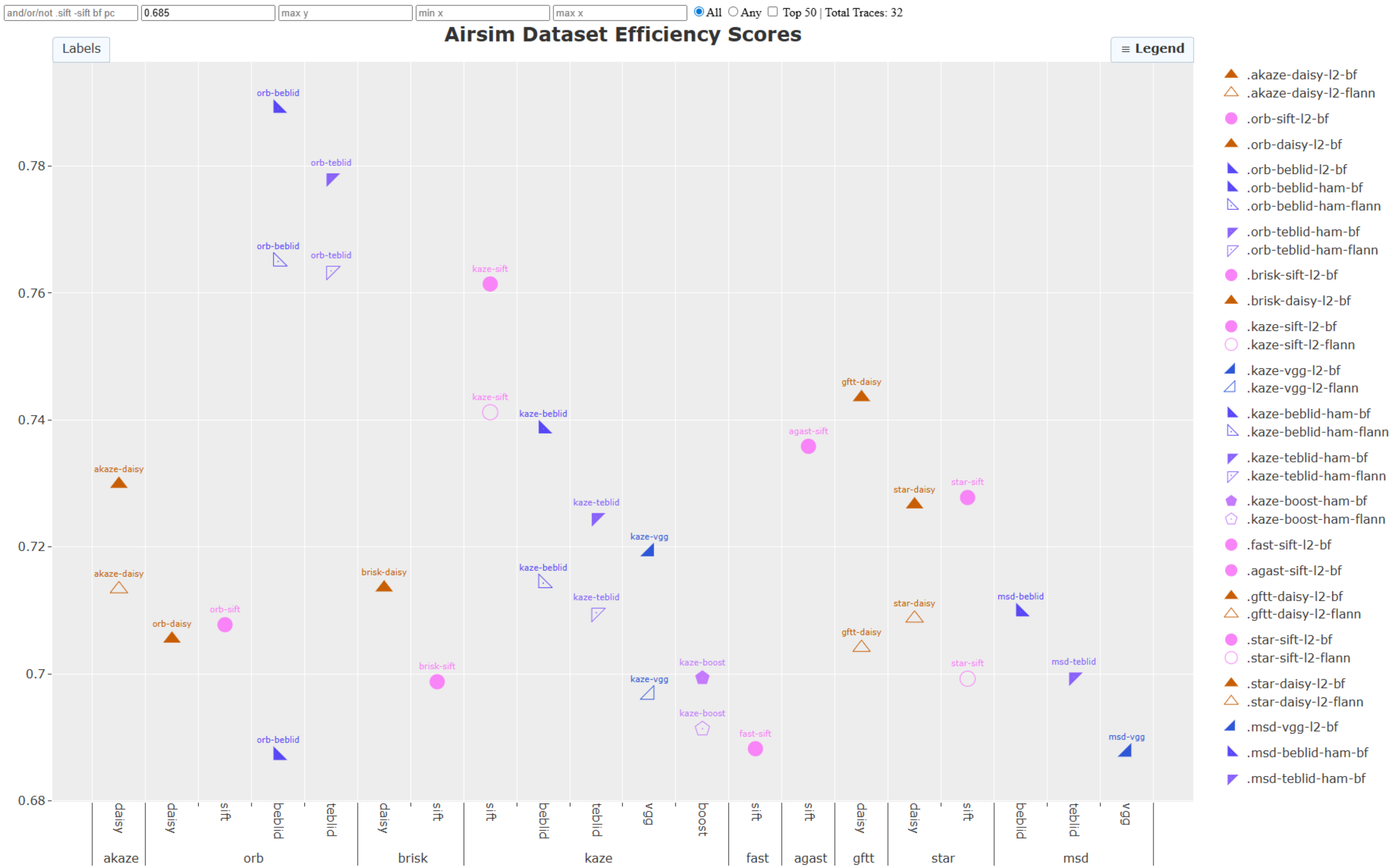


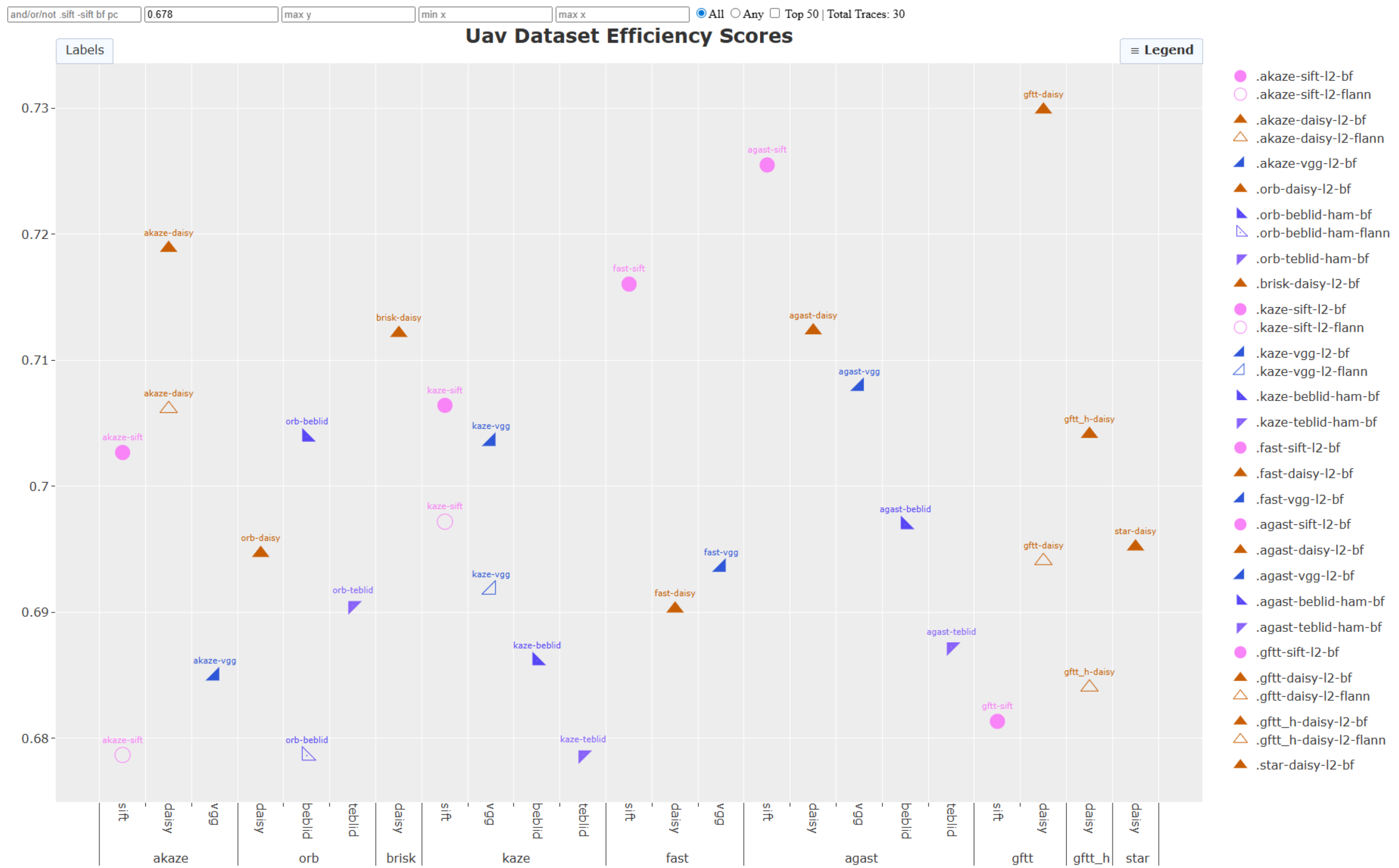
and/or/not .sift -sift bf pc
min y
0.176
min x
max x
All
Any
Top 50 | Total Traces: 32
Drone Dataset Efficiency Scores
Labels
≡ Legend
0.175
0.17
0.165
0.16
0.155
0.15
0.145
latch
sift
latch
akaze
latch
orb
brisk
freak
brief
gftt
latch
brief
gftt_h
latch
star
latch
brisk
freak
brief
hl
latch
brief
msd
latch
orb
brisk
freak
brief
tbmr
.sift-latch-ham-flann
.akaze-latch-l2-bf
.gftt-orb-ham-bf
.gftt-orb-ham-flann
.gftt-brisk-ham-flann
.gftt-freak-ham-bf
.gftt-freak-ham-flann
.gftt-brief-ham-bf
.gftt-brief-ham-flann
.gftt-latch-ham-bf
.gftt_h-brief-ham-flann
.gftt_h-latch-ham-bf
.star-latch-l2-bf
.star-latch-ham-bf
.star-latch-ham-flann
.hl-brisk-ham-bf
.hl-freak-ham-bf
.hl-freak-ham-flann
.hl-brief-ham-flann
.hl-latch-ham-flann
.msd-brief-ham-bf
.msd-brief-ham-flann
.msd-latch-l2-bf
.msd-latch-ham-bf
.msd-latch-ham-flann
.tbmr-orb-ham-bf
.tbmr-brisk-ham-bf
.tbmr-freak-ham-bf
.tbmr-freak-ham-flann
.tbmr-brief-ham-bf
.tbmr-latch-ham-bf
.tbmr-latch-ham-flann

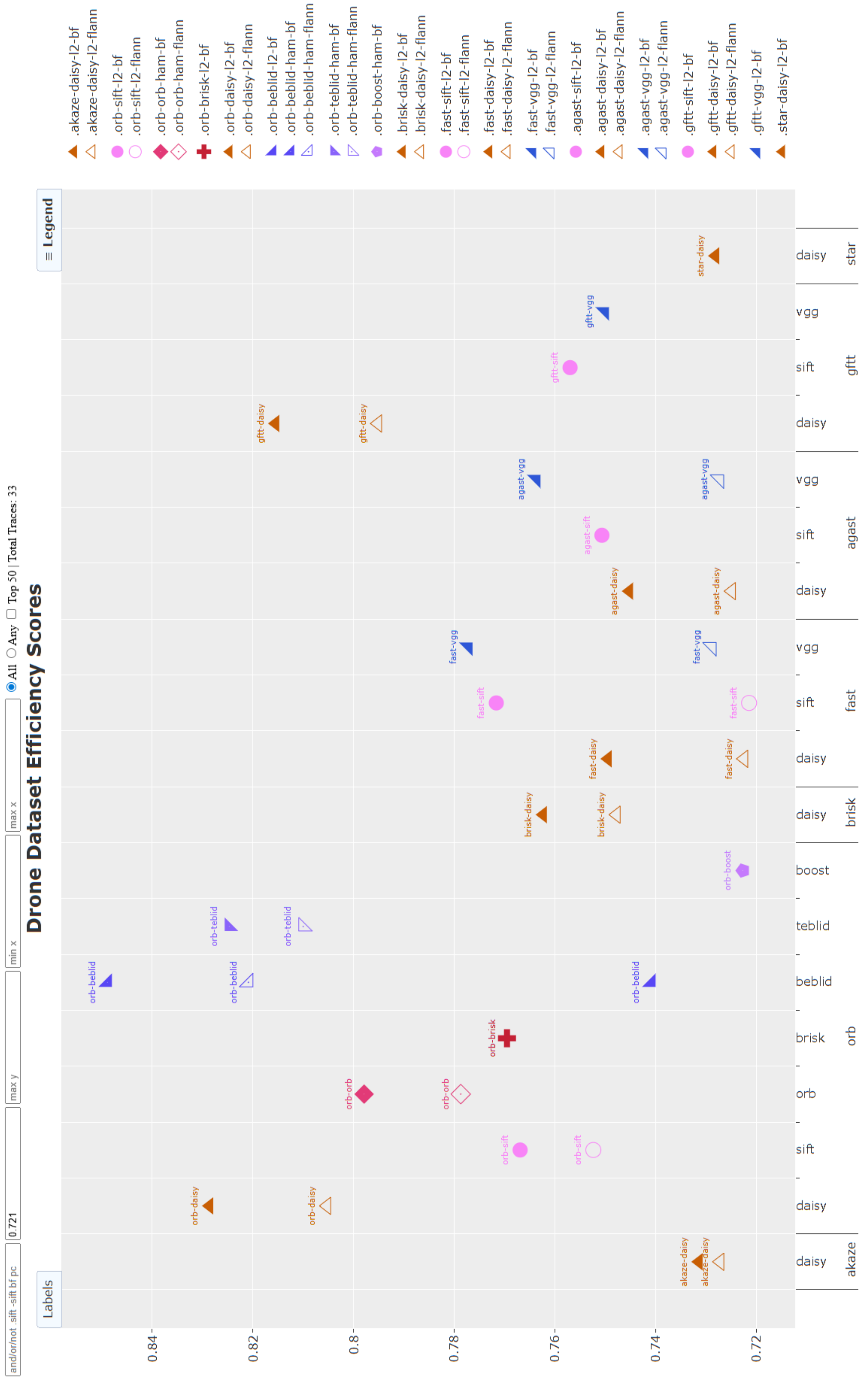
Appendix D.2: Best Efficiency Scores




Oxford Dataset Efficiency Scores
and/or/not sift -sift bf pc
0.5504
min x
max x
max y
All
Any
Top 50 | Total Traces: 32
Labels
≡ Legend
.akaze-sift-l2-bf
.akaze-sift-l2-flann
.akaze-akaze-ham-bf
.akaze-akaze-ham-flann
.akaze-vgg-l2-bf
.akaze-beblid-ham-bf
.akaze-beblid-ham-flann
.akaze-teblid-ham-bf
.akaze-teblid-ham-flann
.akaze-boost-ham-bf
.orb-sift-l2-bf
.orb-orb-ham-bf
.orb-orb-ham-flann
.orb-beblid-l2-bf
.orb-beblid-ham-bf
.orb-beblid-ham-flann
.orb-teblid-ham-bf
.orb-teblid-ham-flann
.brisk-sift-l2-bf
.brisk-beblid-ham-bf
.brisk-teblid-ham-bf
.kaze-sift-l2-bf
.kaze-sift-l2-flann
.kaze-kaze-l2-bf
.kaze-beblid-ham-bf
.kaze-beblid-ham-flann
.kaze-teblid-ham-bf
.kaze-teblid-ham-flann
.agast-sift-l2-bf
.agast-beblid-ham-bf
.gftt-sift-l2-bf
.star-sift-l2-bf
0.62
0.61
0.6
0.59
0.58
0.57
0.56
0.55
sift
akaze
vgg
beblid
teblid
boost
sift
beblid
teblid
orb
sift
beblid
teblid
sift
beblid
teblid
kaze
sift
beblid
sift
sift
akaze
orb
brisk
kaze
agast
gftt
star
akaze-sift
akaze-akaze
akaze-vgg
akaze-beblid
akaze-teblid
akaze-boost
orb-sift
orb-beblid
orb-teblid
orb-orb
brisk-sift
brisk-beblid
brisk-teblid
kaze-sift
kaze-beblid
kaze-teblid
kaze-kaze
agast-sift
agast-beblid
gftt-sift
star-sift

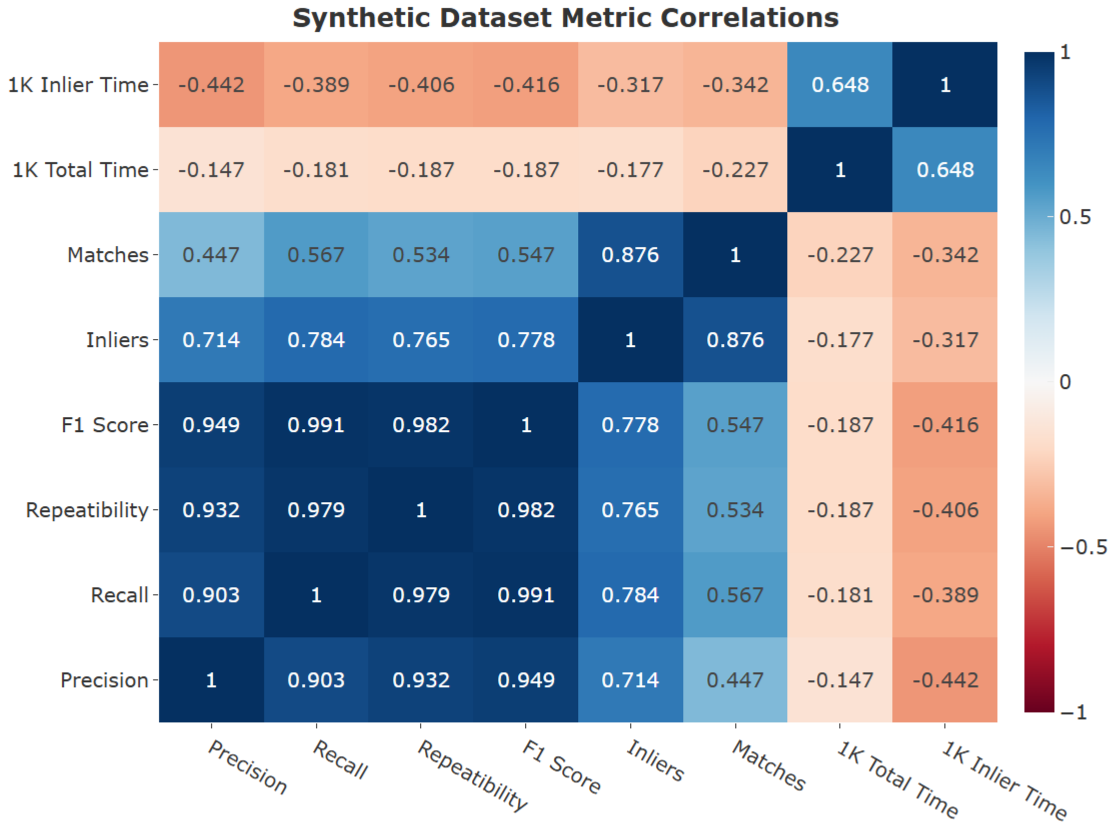

and/or/not .sift -sift bf pc
0.685
max y
min x
max x
All Any Top 50 | Total Traces: 32
Airsim Dataset Efficiency Scores
Labels
≡ Legend
0.78
0.76
0.74
0.72
0.7
0.68
akaze-daisy
orb-daisy
orb-sift
orb-beblid
orb-teblid
brisk-daisy
brisk-sift
kaze-sift
kaze-beblid
kaze-teblid
kaze-vgg
kaze-boost
fast-sift
agast-sift
gftt-daisy
star-daisy
star-sift
msd-beblid
msd-teblid
msd-vgg
daisy
sift
beblid
teblid
vgg
boost
akaze
orb
brisk
kaze
fast
agast
gftt
star
msd
.akaze-daisy-l2-bf
.akaze-daisy-l2-flann
.orb-sift-l2-bf
.orb-daisy-l2-bf
.orb-beblid-l2-bf
.orb-beblid-ham-bf
.orb-beblid-ham-flann
.orb-teblid-ham-bf
.orb-teblid-ham-flann
.brisk-sift-l2-bf
.brisk-daisy-l2-bf
.kaze-sift-l2-bf
.kaze-sift-l2-flann
.kaze-vgg-l2-bf
.kaze-vgg-l2-flann
.kaze-beblid-ham-bf
.kaze-beblid-ham-flann
.kaze-teblid-ham-bf
.kaze-teblid-ham-flann
.kaze-boost-ham-bf
.kaze-boost-ham-flann
.fast-sift-l2-bf
.agast-sift-l2-bf
.gftt-daisy-l2-bf
.gftt-daisy-l2-flann
.star-sift-l2-bf
.star-sift-l2-flann
.star-daisy-l2-bf
.star-daisy-l2-flann
.msd-vgg-l2-bf
.msd-beblid-ham-bf
.msd-teblid-ham-bf

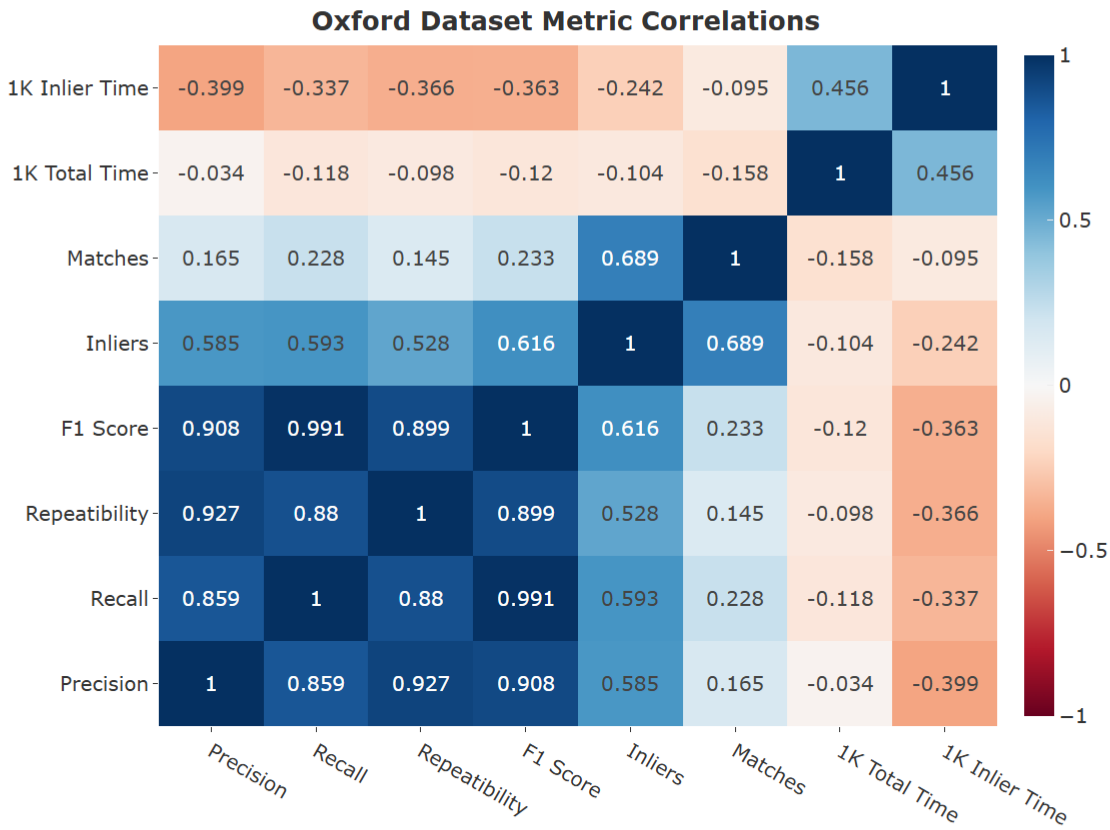

and/or/not .sift -sift bf pc
0.678
max y
min x
max x
All
Any
Top 50 | Total Traces: 30
Uav Dataset Efficiency Scores
Labels
≡ Legend
0.73
0.72
0.71
0.7
0.69
0.68
gftt-daisy
agast-sift
akaze-daisy
fast-sift
brisk-daisy
agast-daisy
agast-vgg
kaze-sift
akaze-daisy
orb-beblid
kaze-vgg
gftt_h-daisy
akaze-sift
kaze-sift
agast-beblid
orb-daisy
star-daisy
gftt-daisy
fast-vgg
kaze-vgg
orb-teblid
fast-daisy
agast-teblid
kaze-beblid
akaze-vgg
gftt_h-daisy
gftt-sift
kaze-teblid
akaze-sift
orb-beblid
sift
daisy
vgg
daisy
beblid
teblid
daisy
sift
vgg
beblid
teblid
sift
daisy
vgg
sift
daisy
vgg
beblid
teblid
sift
daisy
daisy
daisy
akaze
orb
brisk
kaze
fast
agast
gftt
gftt_h
star
.akaze-sift-l2-bf
.akaze-sift-l2-flann
.akaze-daisy-l2-bf
.akaze-daisy-l2-flann
.akaze-vgg-l2-bf
.orb-daisy-l2-bf
.orb-beblid-ham-bf
.orb-beblid-ham-flann
.orb-teblid-ham-bf
.brisk-daisy-l2-bf
.kaze-sift-l2-bf
.kaze-sift-l2-flann
.kaze-vgg-l2-bf
.kaze-vgg-l2-flann
.kaze-beblid-ham-bf
.kaze-teblid-ham-bf
.fast-sift-l2-bf
.fast-daisy-l2-bf
.fast-vgg-l2-bf
.agast-sift-l2-bf
.agast-daisy-l2-bf
.agast-vgg-l2-bf
.agast-beblid-ham-bf
.agast-teblid-ham-bf
.gftt-sift-l2-bf
.gftt-daisy-l2-bf
.gftt-daisy-l2-flann
.gftt_h-daisy-l2-bf
.gftt_h-daisy-l2-flann
.star-daisy-l2-bf

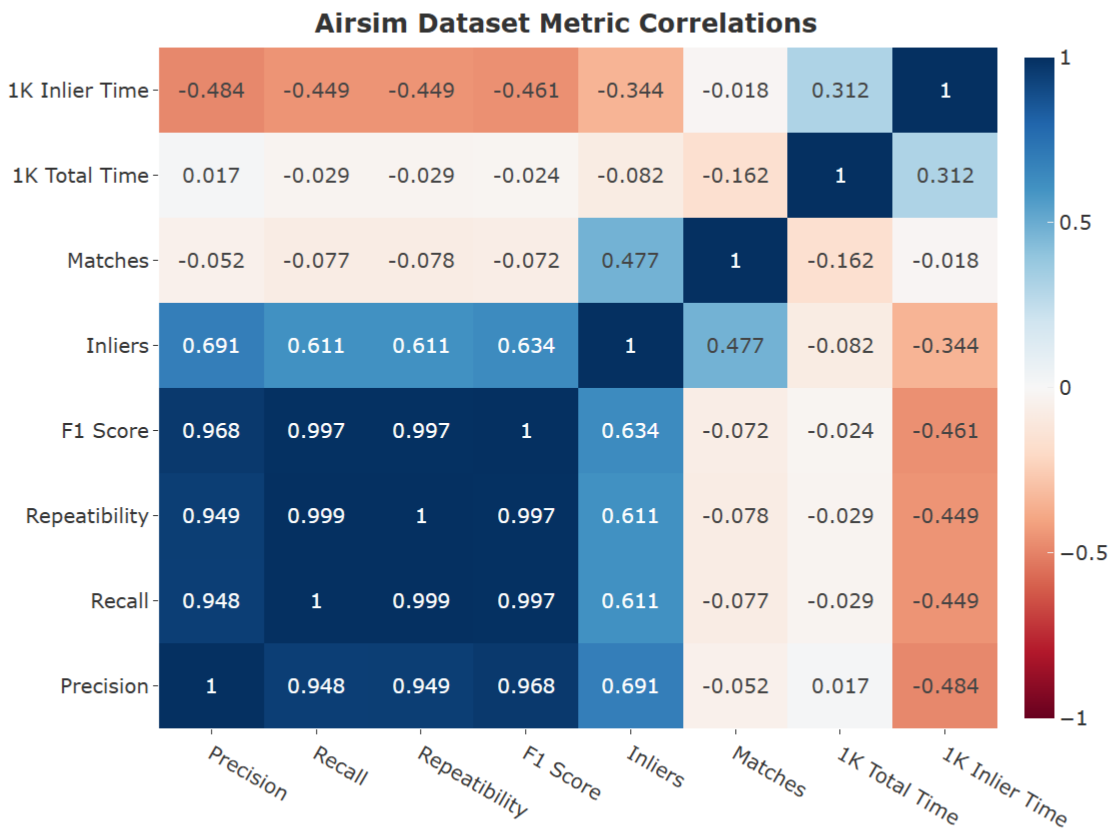


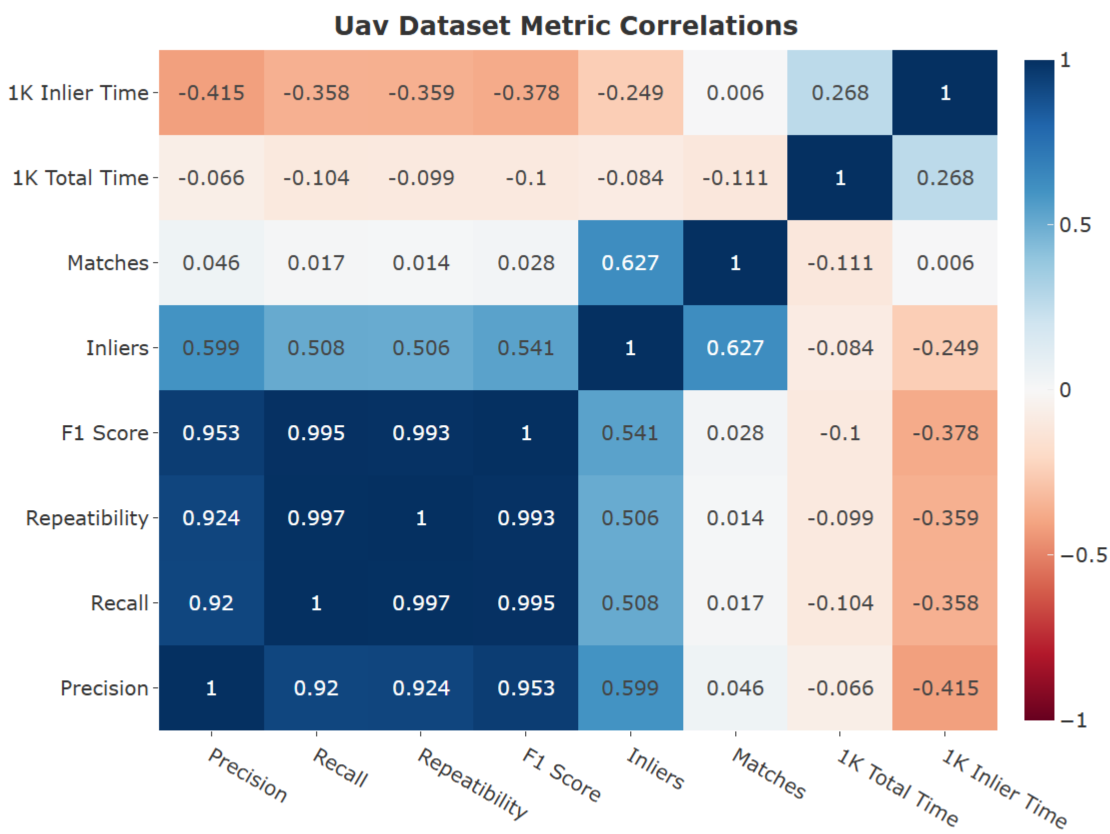
Drone Dataset Efficiency Scores
and/or/not sift -sift bf pc
0.721
max y
min x
max x
All Any Top 50 | Total Traces: 33
Labels
≡ Legend
0.84
0.82
0.8
0.78
0.76
0.74
0.72
akaze: daisy
orb: daisy, sift, orb, brisk, beblid, teblid, boost
brisk: daisy
fast: daisy, sift, vgg
agast: daisy, sift, vgg
gftt: daisy, sift, vgg
star: daisy
.akaze-daisy-l2-bf
.akaze-daisy-l2-flann
.orb-sift-l2-bf
.orb-sift-l2-flann
.orb-orb-ham-bf
.orb-orb-ham-flann
.orb-brisk-l2-bf
.orb-daisy-l2-bf
.orb-daisy-l2-flann
.orb-beblid-l2-bf
.orb-beblid-ham-bf
.orb-beblid-ham-flann
.orb-teblid-ham-bf
.orb-teblid-ham-flann
.orb-boost-ham-bf
.brisk-daisy-l2-bf
.brisk-daisy-l2-flann
.fast-sift-l2-bf
.fast-sift-l2-flann
.fast-daisy-l2-bf
.fast-daisy-l2-flann
.fast-vgg-l2-bf
.fast-vgg-l2-flann
.agast-sift-l2-bf
.agast-daisy-l2-bf
.agast-daisy-l2-flann
.agast-vgg-l2-bf
.agast-vgg-l2-flann
.gftt-sift-l2-bf
.gftt-daisy-l2-bf
.gftt-daisy-l2-flann
.gftt-vgg-l2-bf
.star-daisy-l2-bf

Appendix E: Metrics Correlation Matrices

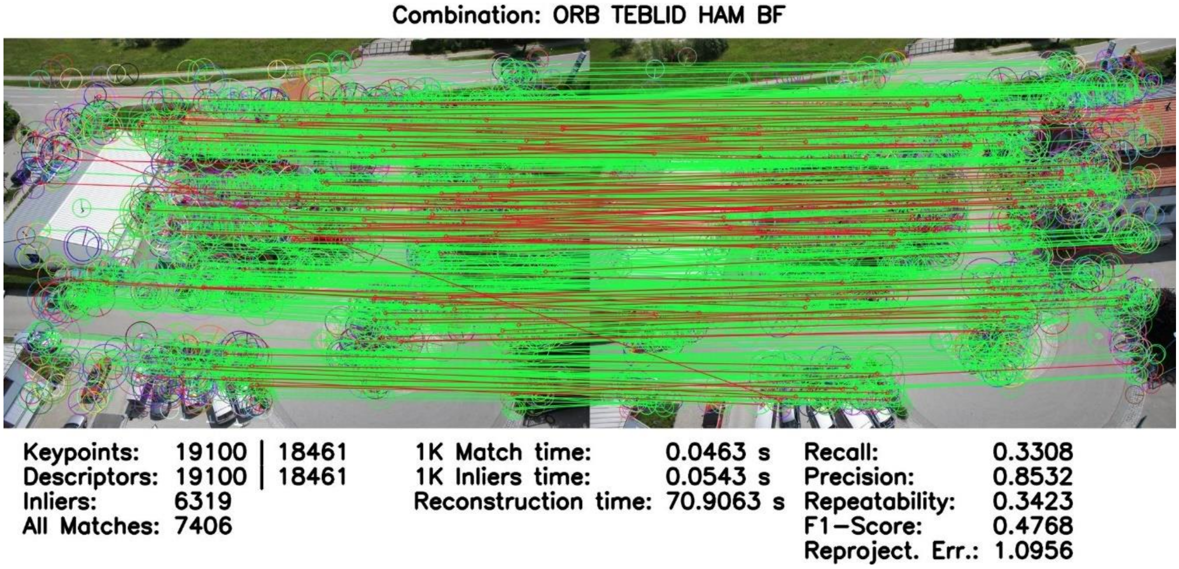


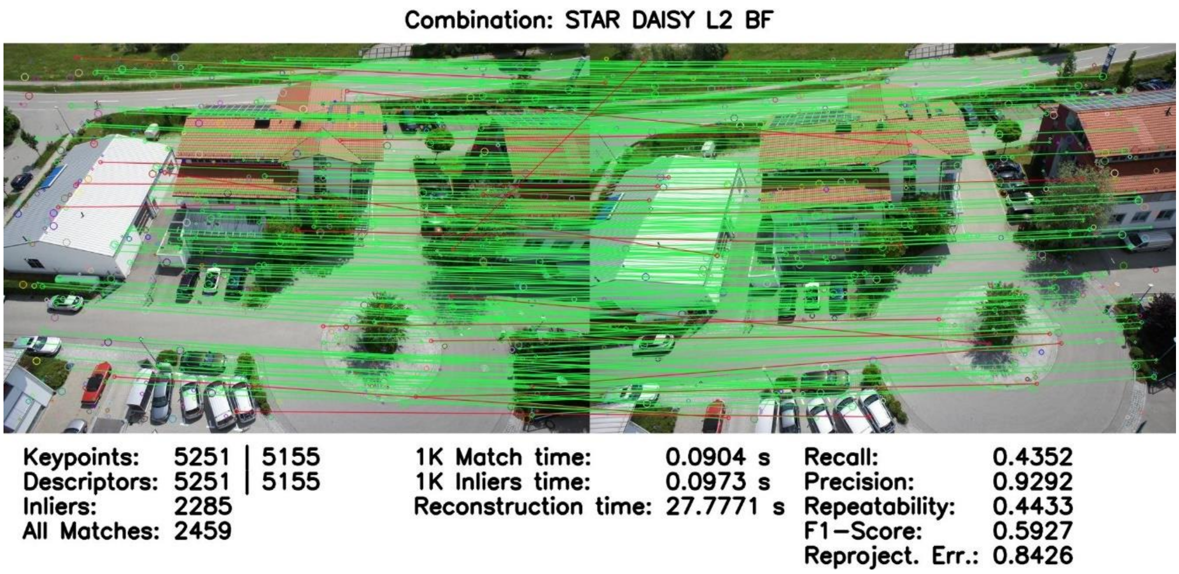



Airsim Dataset Metric Correlations
1K Inlier Time -0.484 -0.449 -0.449 -0.461 -0.344 -0.018 0.312 1
1K Total Time 0.017 -0.029 -0.029 -0.024 -0.082 -0.162 1 0.312
Matches -0.052 -0.077 -0.078 -0.072 0.477 1 -0.162 -0.018
Inliers 0.691 0.611 0.611 0.634 1 0.477 -0.082 -0.344
F1 Score 0.968 0.997 0.997 1 0.634 -0.072 -0.024 -0.461
Repeatibility 0.949 0.999 1 0.997 0.611 -0.078 -0.029 -0.449
Recall 0.948 1 0.999 0.997 0.611 -0.077 -0.029 -0.449
Precision 1 0.948 0.949 0.968 0.691 -0.052 0.017 -0.484
Precision
Recall
Repeatibility
F1 Score
Inliers
Matches
1K Total Time
1K Inlier Time
1
0.5
0
−0.5
−1

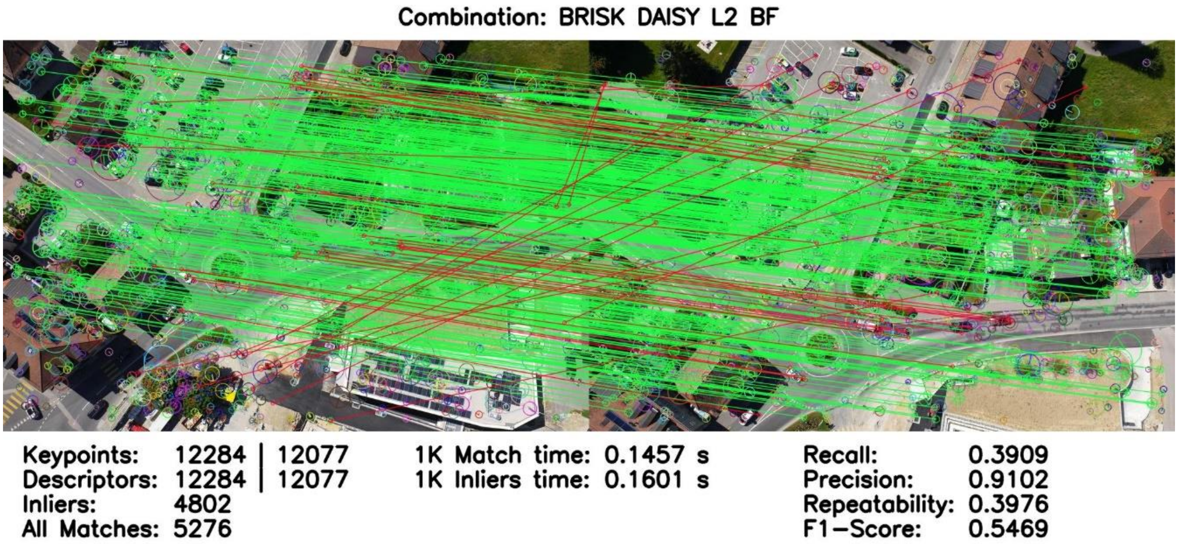

Uav Dataset Metric Correlations
1K Inlier Time -0.415 -0.358 -0.359 -0.378 -0.249 0.006 0.268 1
1K Total Time -0.066 -0.104 -0.099 -0.1 -0.084 -0.111 1 0.268
Matches 0.046 0.017 0.014 0.028 0.627 1 -0.111 0.006
Inliers 0.599 0.508 0.506 0.541 1 0.627 -0.084 -0.249
F1 Score 0.953 0.995 0.993 1 0.541 0.028 -0.1 -0.378
Repeatibility 0.924 0.997 1 0.993 0.506 0.014 -0.099 -0.359
Recall 0.92 1 0.997 0.995 0.508 0.017 -0.104 -0.358
Precision 1 0.92 0.924 0.953 0.599 0.046 -0.066 -0.415
Precision
Recall
Repeatibility
F1 Score
Inliers
Matches
1K Total Time
1K Inlier Time
1
0.5
0
−0.5
−1

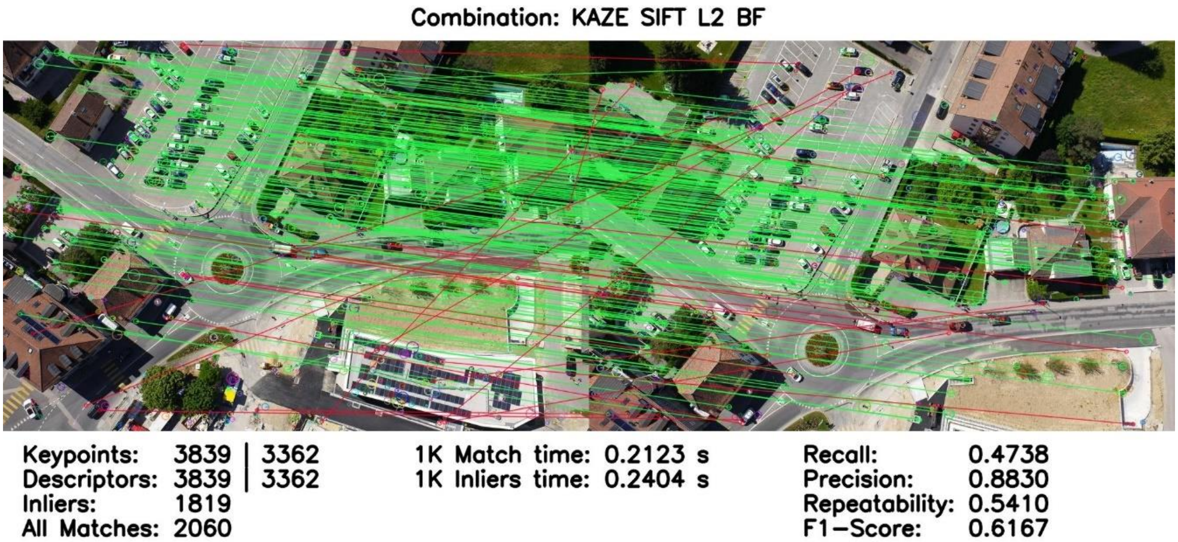
Appendix F.1: Drone Dataset: Binary vs Floating-Point Descriptors

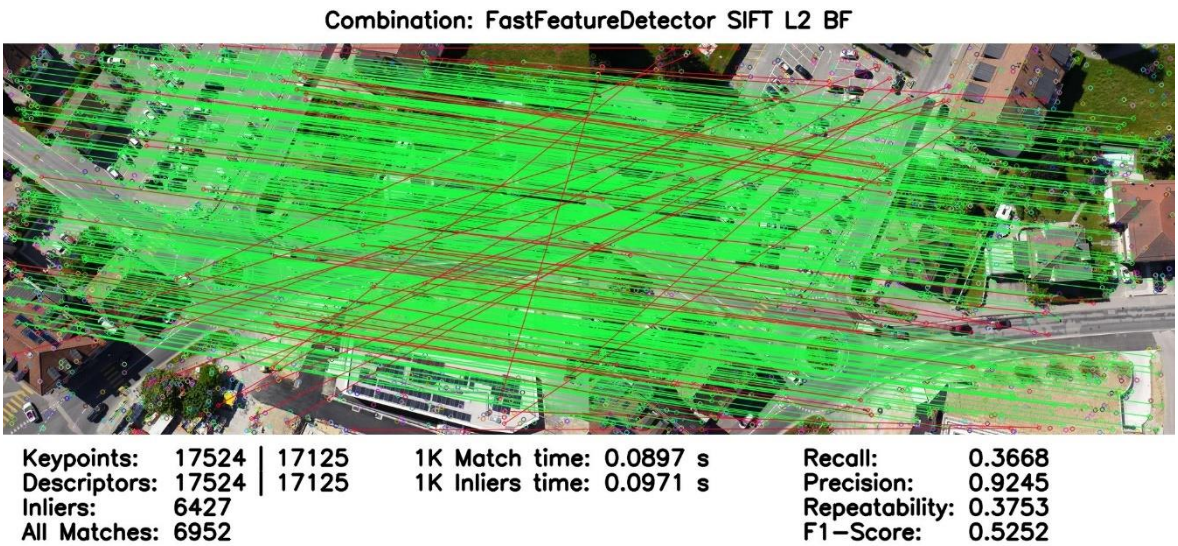




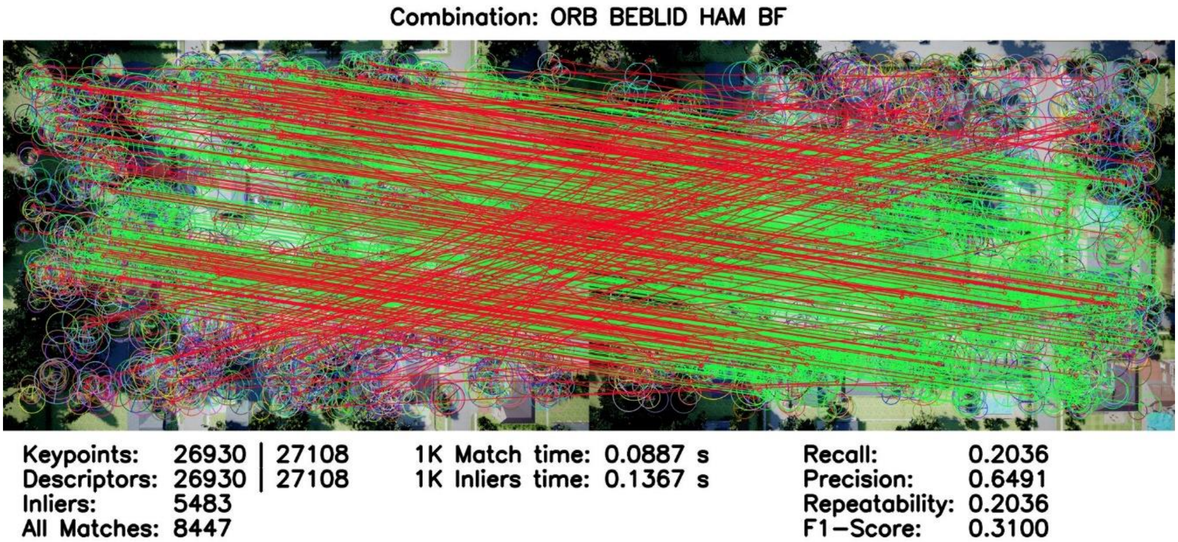
Appendix F.2: Diverse UAV Scenarios: Detector and Descriptor Alternatives

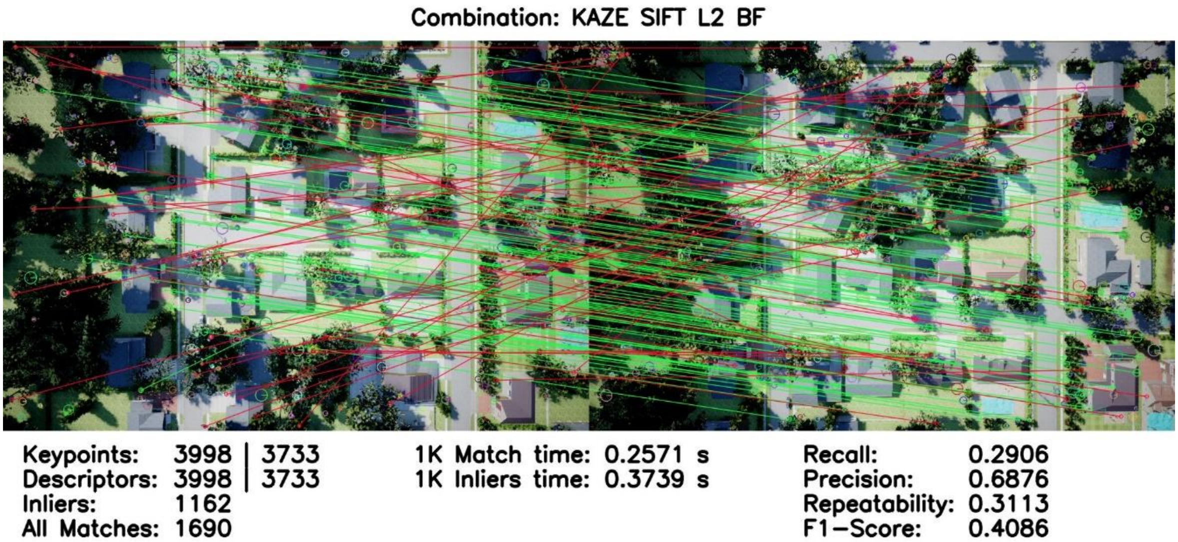


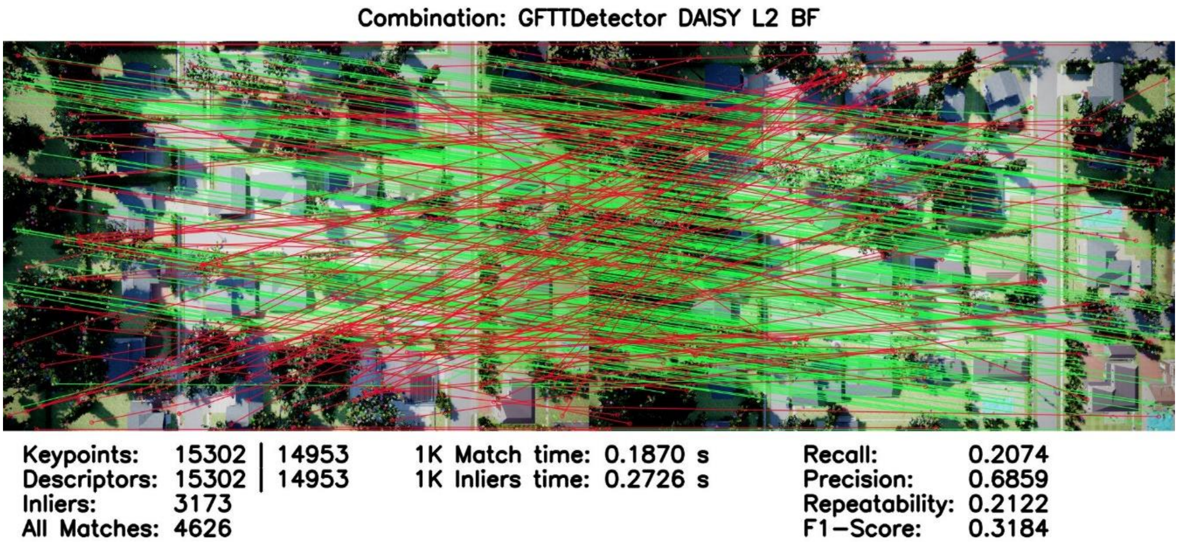




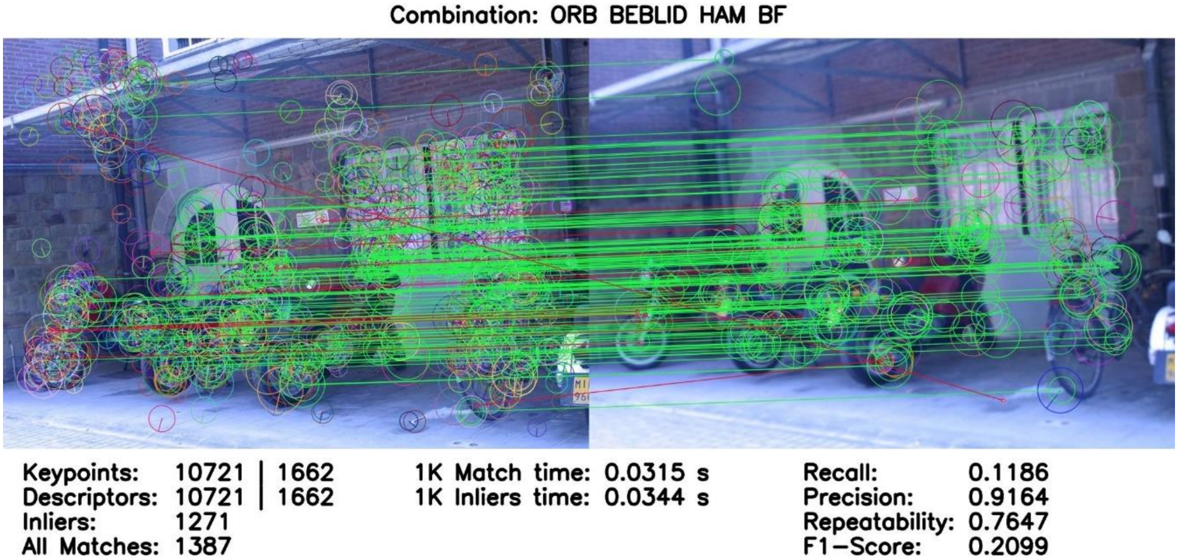
Appendix F.3: AirSim Controlled Environment: Efficiency vs Consistency Trade-off

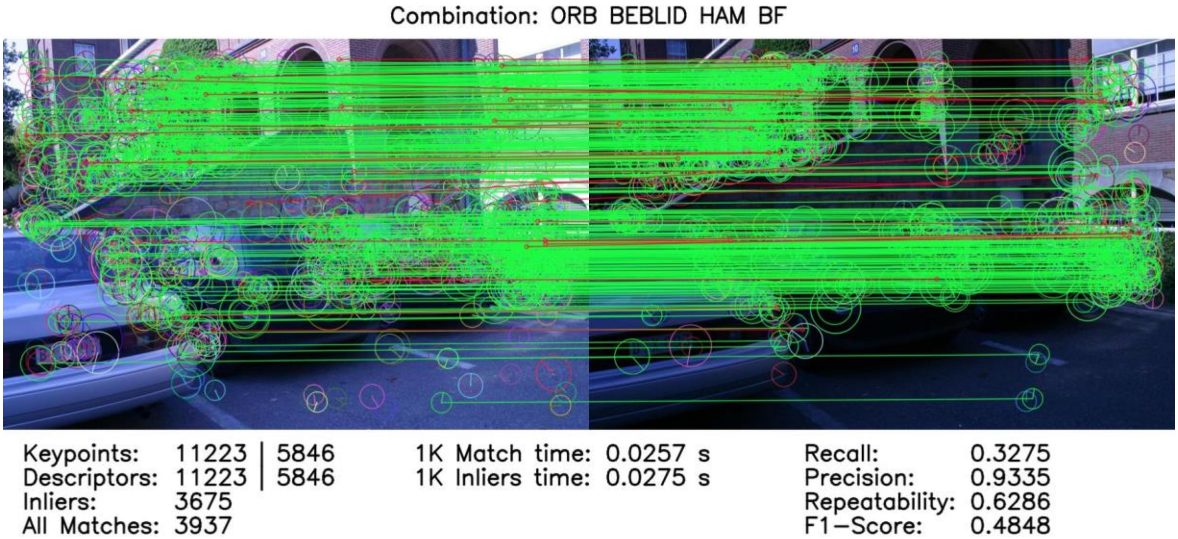


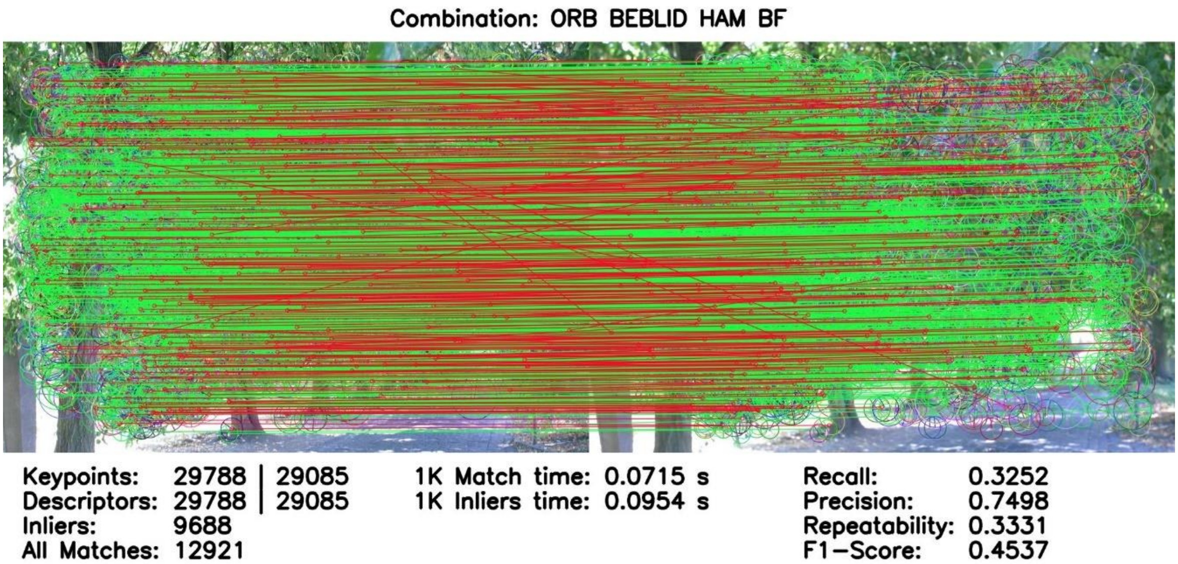




Appendix F.4: Oxford Benchmark Robustness: ORB-BEBLID







Appendix F.5: Synthetic Dataset Evaluation: ORB-ORB Peak Performance