

# Üç Boyutlu Sahneler ve Nesne Tanıma için Gürbüz Anahtar Nokta Eşleştirilmesi Robust Keypoint Matching for Three Dimensional Scenes and Object Recognition

Ali Köksal, Furkan Eren Uzyıldırım, Mustafa Özuysal  
Bilgisayar Mühendisliği Bölümü  
İzmir Yüksek Teknoloji Enstitüsü, İzmir, Türkiye  
{alikoksal,furkanuzyildirim,mustafaozuysal}@iyte.edu.tr

**Özetçe** —Bu çalışmada yakın zamanda düzlemsel nesnelere için ikilik betimleyiciler ile anahtar nokta eşleme amacıyla önerilen bir yöntem üç boyutlu nesnelere için uyarlanmıştır. Bu yöntemin başarısı yüzden fazla resim içeren bir müze nesne tanıma uygulamasında test edilmiştir. Ayrıca sadece eşleşme başarısı yüksek betimleyicilerin kullanılmasının nesne tanıma uygulamasının başarısına etkisi de ölçülmüştür.

**Anahtar Kelimeler**—Bilgisayarla görü, nesne tanıma, ikilik betimleyiciler, anahtar nokta eşleme

**Abstract**—In this paper, we adapt a recently proposed keypoint matching approach for binary descriptors and planar objects to three dimensional objects. We also evaluate the performance of this approach for a museum object recognition application containing more than one hundred paintings. Moreover, we quantify the effect of selecting only descriptors with high matching ratio on the success rate of the object recognition application.

**Keywords**—Computer vision, object recognition, binary descriptors, keypoint matching

## I. GİRİŞ

Nesne tanıma problemi daha önceden bilinen ve sınırlı sayıda bakış açısından görüntüleri toplanmış nesnelere içeriği bilinmeyen bir sorgu görüntü içerisinde yer alıp almadığının tespitidir. Bu problemin çözümü için yaygın olarak sorgu görüntü ile daha önceden çekilmiş görüntüler eşleştirilir. Genelde, bu eşleştirme görüntülerden tespit edilen anahtar noktalar [1], [2] üzerinden gerçekleştirilir. Her bir anahtar nokta etrafındaki dokuyu tanımlayan bir betimleyici [3], [4] hesaplanır ve birbirine yakın betimleyicilere karşılık gelen anahtar noktaların doğru eşler olduğu varsayılır.

Hesaplama gücünün sınırlı olduğu durumlarda genellikle ikili betimleyiciler [4]–[9] tercih edilmektedir. Bu betimleyicilerin hem hesaplanması kolaydır hem de iki betimleyici arasındaki uzaklık sadece bir kaç makine komutu ile hesaplanabilir.

Ancak bu betimleyiciler görüntünün elde edildiği bakış açısından SIFT [10] benzeri betimleyicilere göre daha çok etkilenirler. Burada temel etmen bu betimleyicilerin ikilik sayı sistemini kullanmasıdır. Sonuç olarak bu betimleyiciler kullanılarak yapılan anahtar nokta eşleşmesinde görece daha çok hata yapılır. Çünkü sorgu görüntüdeki anahtar noktanın betimleyicisinin, ait olduğu nesne bilinen betimleyiciler arasındaki en yakın komşusu doğru nesneye ait olmayabilir. Aslında büyük nesne kümeleri ile çalışırken en yakın komşu eşsiz dahi olmayabilir. Bu durumda da eşleme hataları ortaya çıkar ve nesne tanıma başarısını düşer.

Yakın zamanda bu durumu önlemek için en yakın tek bir komşu yerine, en yakın  $K$  komşunun hesaplanmasını sağlayan bir yöntem önerilmiştir [11]. Bu yöntemde, sorgu görüntüde yer alan bir betimleyiciye olan uzaklıklara göre sıralama yapılarak en yakın  $K$  betimleyici hesaplanmakta ve doğru eş betimleyici bu  $K$  elemanlı küme içinde olasılığa dayalı bir yöntem ile seçilmektedir. Seçimi etkileyen olasılık değerleri, farklı bakış açılarından betimleyicilerdeki değişimleri modellemektedir. Yöntem sadece düzlemsel nesnelere ile test edildiğinden olasılık değerleri sentetik olarak oluşturulan eğitim görüntülerinden hesaplanmaktadır.

Üç boyutlu nesnelere farklı bakış açılarından görüntülerini sentetik olarak oluşturmak düzlemsel nesnelere kadar basit değildir. Bu nedenle, eğitim amacıyla kullanılacak verilerin farklı açılardan çekilmiş görüntülerden toplanması daha uygundur. Bu çalışmada [11] tarafından önerilen yöntem, üç boyutlu nesnelere için uyarlanmış ve bu nesnelere üzerindeki anahtar noktaların eşleşmesi için başarısı test edilmiştir.

Aynı yöntem, çekilen fotoğrafların bir müzedeki yaklaşık yüz farklı eser içerisinden hangisine ait olduğunun belirlendiği bir nesne tanıma uygulamasına da uyarlanmıştır. Yapılan deneylerin sonuçları, nesne tanıma başarısının ilk on yakın komşu kullanılarak artırılabilirliğini ve betimleyicilerin sadece yarısı kullanılsa bile eserlerin tanınabilirliğini göstermektedir.

## II. İLGİLİ LİTERATÜR

Çalışmada üç boyutlu sahnelerde anahtar nokta eşleme ve nesne tanıma için [11] tarafından önerilen yöntem kullanılıp test edilmiştir. Bu yöntemle en yakın yaklaşım [12] tarafından

Bu bildiriye yer alan çalışmalar TÜBİTAK tarafından 113E496 numaralı araştırma projesi kapsamında desteklenmiştir.

önerilen ve BOLD adı verilen maskeleye yaklaşımdır. BOLD betimleyicisi bir eğitim verisinden betimleyici değişimlerini modellemek yerine betimleyicinin güvenilir ikilik değerlerini anahtar nokta bazında öğrenir. Her bir anahtar nokta için bu ikilik değerleri içeren bir maske saklanır ve betimleyici eşleşmesinde kullanılır. Üç boyutlu nesnelere için BOLD betimleyicisi hesaplanması ve test edilmesi için de önerilen yöntem kullanılabilir.

Anahtar nokta betimleme amacıyla çalışmada BRIEF betimleyicisi kullanılmış ancak buna alternatif olarak [7]–[9] tarafından geliştirilen betimleyiciler de kullanılabilir. [11] tarafından önerilen yöntem bu betimleyiciler ile de uyumlu olduğundan çalışmada kullanılan teknikler doğrudan bu betimleyiciler için de uyarlanabilir.

Anahtar nokta veri kümelerinin sıkıştırılması konusunda robotik alanında çalışmalar mevcuttur. Bir ortamda bulunan tüm anahtar noktaların saklanması gerçek zamanlı robotik uygulamaları için de ortam büyüklüğü arttıkça sorun olmaktadır. [13] nokta kümelerinin azaltılması için olasılıksal bir yaklaşım geliştirmiştir. [14] ise nokta kümeleri ve sahneler üzerinden bir tam sayı eniyileme problemi çözümüyle noktalar arasındaki ilişkiyi de kullanarak bazı noktaları eler. Çalışmada kullanılan yaklaşım bu yaklaşımlardan daha basit ama istatistiklerinin hesaplanması için kullanılan eğitim aşamasıyla daha uyumludur.

### III. YÖNTEM

İkili betimleyiciler ayrık yapıları sebebiyle görüntünün elde edildiği bakış açısından [10] benzeri betimleyicilere göre daha çok etkilenirler ve bu durum eşleşme hatalarının ortaya çıkmasına ve nesne tanıma başarısının düşmesine sebep olabilir. Çünkü sorgu betimleyicinin Hamming uzaklığına göre en yakın komşusu doğru nesneye ait olmayabilir ve hatta büyük nesne kümeleri için en yakın komşu eşsiz dahi olmayabilir.

Bu çalışmada, ikili betimleyiciler kullanılarak üç boyutlu nesne tanıma için önce her bir sorgu betimleyiciye ait en yakın komşu yerine  $K$  adet en yakın komşu belirlenip bundan sonra bu listedeki aday anahtar noktalar için özel bir betimleme yapan iki kademeli bir yöntem kullanılmıştır [11]. Bu yöntemde, ilk adımda her bir sorgu betimleyici için bilinen nesnelere ait betimleyicilerden Hamming uzaklığına göre en yakın  $K$  tanesi bulunarak bir listede saklanır. İkinci adımda ise bu betimleyicilerden en uygunu, olasılığa dayanarak ve her bir anahtar nokta için özel olarak hesaplanan bir değer yardımıyla, seçilir. Bu değer hesaplanması için betimleyicilerdeki değişimleri modelleyen ve önceden bilinen görüntülerden hesaplanmış olan veriden faydalanılır. Kullanılan yöntem, düzlemsel nesnelere için anahtar nokta eşleşmesinde başarı sağlamıştır. Ancak yöntem üç boyutlu nesnelere ile test edilmemiş ve nesne tanıma başarısı ayrıca ölçülmemiştir.

#### A. Üç Boyutlu Nesnelere için Eğitim Kümesinin Hesaplanması

Kullanılan yöntemde betimleyicilerin kamera bakış açısına bağlı değişimleri istatistiksel olarak modellenirken ilgin dönüşümler ile sentezlenen eğitim verisinden yararlanılmıştır. Bilinen nesnelere düzlemsel olduğunda bu gerçekçi bir yaklaşımdır. Ancak nesnelere üç boyutlu olduğunda tek bir görüntüden eğitim verisi çıkarılması gerçekçi değildir. Bu çalışmada, bilinen üç boyutlu görüntülerden, çevrimdışı bir eğitim aşamasında, veri çıkarmak için her bir üç boyutlu nesnenin farklı kamera



Şekil 1: Üç boyutlu bir nesnenin çevrimdışı eğitim aşamasında kullanılan farklı kamera pozisyonlarından elde edilmiş görüntü örnekleri. (a) Üç boyutlu nesnenin referans görüntüsü. (b) Üç boyutlu nesnenin kamera ekseninde kırk derece döndürülmüş görüntüsü. (c) Üç boyutlu nesnenin kamera ekseninde doksan derece döndürülmüş görüntüsü.



Şekil 2: Çalışmada kullanılan üç farklı üç boyutlu nesneye ait veri kümelerinin referans görüntüleri.

pozisyonlarındaki görüntülerinden oluşan veri kümesi [15] kullanılmıştır.

Bu veri kümesi dönen bir tablaya yerleştirilmiş yüzden fazla nesneyi barındırmaktadır. Şekil 2’de örnek nesnelere gösterilmiştir. Şekil 1’de nesnelere çevrimdışı eğitim aşamasında kullanılan veri kümesinden örnekler gösterilmiştir. Veri kümesi dönen tablaya konulan nesnelere görüntüleri alt ve üst kamera olmak üzere iki kamera tarafından beş derecelik döndürme açılarında çekilmiştir. Veri kümesi nesne görüntülerinin yanında bir damalı düzlemsel kalibrasyon nesnesinin de görüntülerini içermektedir. Bu sayede kameraların iç ve dış kalibrasyon değişkenleri hesaplanabilmiş ve anahtar noktalar üç boyutlu nesnelere farklı görüntülerine taşınabilmiştir. Anahtar noktaların taşınması için iki görüntü arasındaki geometrik ilişkinin tanımlanması gereklidir. Bu tanımlamaya temel matris (fundamental matrix) denilip, bu matris iç ve dış kalibrasyon değişkenleri kullanılarak aşağıdaki formül ile hesaplanır.

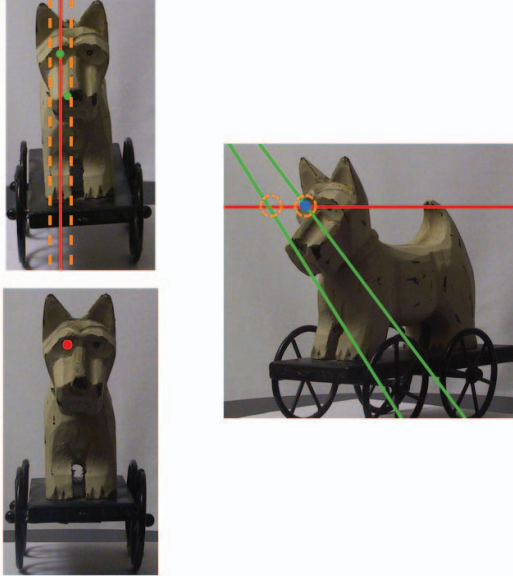
$$F = [e_2]_x P_2 P_1^T \quad (1)$$

Bu formüldeki  $e_2$  ikinci kameranın epipol’ü olup Formül 2 ile hesaplanır.

$$e_2 = P_2 \begin{bmatrix} R_1^T & t_1 \\ 1 & \end{bmatrix} \quad (2)$$

Formül 1 ve 2’deki  $P_1$  ve  $P_2$  kameraların izdüşüm matrisleri olup üç boyutlu geri çatım elde edilmesinde kullanılır [16].

Eğitim verisinin hesaplanması amacıyla, aynı üç boyutlu noktanın farklı bakış açılarından iki boyutlu konumunu bulmak için görüntülerdeki anahtar noktalar geometrik kıstaslar yardımıyla eşleştirilmiştir [15]. Bunun için görüntülerdeki anahtar noktalar diğer görüntülere taşınmıştır. Nesnelere üç boyutlu olduğu için bir görüntüdeki konumlar başka bir görüntüye doğru parçası olarak taşınabilmektedir. Ancak bir görüntüdeki konumun diğer görüntüdeki eşleşmesini bulmak için nesnenin üç görüntüsünün kullanılması gerekmektedir. Testlerde sıfır



Şekil 3: Sol alttaki resim *Dog* nesnesinin sıfır derecede alt kamera tarafından ( $I_{A0}$ ), sol üstteki resim sıfır derece üst kamera tarafından ( $I_{Ü0}$ ) ve sağdaki resim kırk derece alt kamera tarafından ( $I_{A40}$ ) çekilmiş görüntülerinden elde edilmiştir.  $I_{A0}$  görüntüsündeki kırmızı nokta ile gösterilmiş anahtar nokta, görüntüler arasındaki geometrik ilişki kullanılarak  $I_{Ü0}$  ve  $I_{A40}$  görüntülerine epipolar çizgi (kırmızı renkli) olarak taşınmıştır.  $I_{Ü0}$  görüntüsündeki epipolar çizgiye Öklit uzaklığı üç pikselden daha az olan iki anahtar nokta yeşil noktalar ile gösterilmiştir. Bu noktalar  $I_{A40}$  görüntüsüne epipolar çizgi (yeşil renkli) olarak taşınmıştır.  $I_{A40}$  görüntüsünde yeşil epipolar çizgiler ile kırmızı epipolar çizginin kesişim noktaları bulunmuş ve bunlara uzaklığı üç pikselden az olan mavi anahtar nokta  $I_{A0}$  görüntüsündeki kırmızı ile gösterilen anahtar noktanın eşleniği olarak bulunmuştur.

derece alt kamera tarafından çekilmiş nesnelerin görüntüsündeki (referans görüntülerdeki) anahtar nokta konumları yine alt kamera tarafından çekilmiş, nesnenin diğer görüntülerine taşınırken sıfır derece üst kamera tarafından çekilmiş görüntü yardımcı olarak kullanılmıştır. Şekil 3 geometrik kıstaslar ile anahtar noktaların nasıl eşlendiğini özetlemektedir.

#### B. Anahtar Nokta Eşleştirme ile Nesne Tanıma

[11] tarafından geliştirilen yöntemin gerçekçi bir nesne tanıma uygulamasında test edilmesi için bir müzedeki tablolara ait imgelerden oluşan bir veri kümesi oluşturulmuştur. Orsay müzesindeki 112 tablonun ön cepheden çekilmiş fotoğrafları bilinen nesnelere kümesini oluşturmaktadır. Referans imge örnekleri Şekil 4 ile gösterilmektedir.

Daha sonra bu veri kümesinde nesne tanıma başarımını ölçmek için 11 adet imge seçilmiş ve bunlar basılarak farklı arkaplan ve açılar ile test veri kümesi oluşturulmuştur. Aynı test imgesi için farklı test koşulları Şekil 5 ile gösterilmektedir. Test imgelerindeki görüntülerde referans nesne konumları el ile işaretlenerek test ve referans imgeler arasındaki iki boyutlu perspektif dönüşümler hesaplanmıştır. Bu dönüşümler kullanılarak yapılacak eşlemelerin doğruluk ya da yanlışlığı tespit edilebilir.



Şekil 4: Orsay Müzesi veri kümesi referans imge örnekleri



Şekil 5: Orsay Müzesi veri kümesi test koşulları örnekleri

Nesne tanıma başarımının ölçülmesi için her bir test imgesi için betimleyiciler çıkarılarak tüm nesnelere hesaplanan betimleyiciler arasında arama yapılmıştır. Bu aramanın sonuçları hem en yakın komşu için hem de 10 en yakın komşu için yerelliğe duyarlı karma [17] (YDK) yöntemiyle hesaplanmıştır. Bu büyüklükteki veri kümelerinde betimleyicilerin büyük çoğunluğu yanlış eşleşeceğinden doğrudan nesne tanıma yapılması mümkün değildir. Bunun yerine eşlenen betimleyicilerin hangi nesnelere ait olduğuna bakılarak en çok betimleyicinin eşlendiği ilk üç nesne belirlenir. Bu ilk üç nesne içinden her birine ayrı ayrı eşleme yapılarak geometrik sınırlamaları da kullanan gürbüz bir yöntemle (PROSAC [18]) eşlemeler kontrol edilir. En çok eşleme yapılan nesne seçilmiş olur.

## IV. DENEYLER

### A. Üç Boyutlu Nesnelere için Anahtar Nokta Eşleştirme

Yöntem bölümünde belirtildiği gibi tek bir imgeden çıkarılacak eğitim verileri üç boyutlu değişimleri doğru modellemeyeceğinden on derece aralıklarda eğitim imgeleri alınmış bunlar zayıf ilgin dönüşümlerle bükülerek eğitim verisi oluşturulmuş ve bu eğitim verisinden de betimleyici davranışlarının olasılıksal dağılımları hesaplanmıştır. Test için eğitim verisinde yer almayan ve beş derece döndürülmüş imgeden başlayıp on derece aralıklarla devam eden imgeler kullanılmıştır.

Deneylerde hem referans hem de test imgelerinde betimleyiciler hesaplanarak gerçek eşlemeye uygun bir model kullanılmıştır. Test imgelerinde FAST yöntemiyle tespit edilen noktaların betimleyicileri için referans betimleyicilerle eşleme yapılmıştır. Eşleme başarımı, doğru eşleşmelerin sayısının toplam eşleşen nokta sayısına bölünmesiyle ölçülmüştür ve elde edilen değere *tanınma oranı* denir.

Bu çalışmada, iki farklı şekilde tanıma oranı ölçülmüştür. İlk olarak, sorgu betimleyicileri ile referans betimleyicileri



TABLO I: Dog, Horse, ve Oil veri kümeleri için tanınma oranı yüzdeleri.

Veri Kümesi	Dog		Horse		Oil		
	EYK	10-EYK	EYK	10-EYK	EYK	10-EYK	
Döndürme Açısı	-45	2.3	<b>8.0</b>	7.6	<b>13.6</b>	5.5	<b>6.6</b>
	-35	9.4	<b>15.1</b>	11.5	<b>18.1</b>	5.6	<b>7.4</b>
	-25	11.9	<b>16.0</b>	16.7	<b>21.2</b>	9.7	<b>11.7</b>
	-15	15.8	<b>20.1</b>	24.6	<b>27.8</b>	<b>18.3</b>	<b>18.3</b>
	-5	24.0	<b>24.8</b>	44.6	<b>44.8</b>	<b>29.4</b>	29.1
	5	<b>33.7</b>	33.2	<b>44.1</b>	<b>44.1</b>	34.9	<b>36.2</b>
	15	13.2	<b>15.0</b>	29.9	<b>32.1</b>	<b>37.9</b>	37.7
	25	5.6	<b>10.8</b>	15.7	<b>24.9</b>	29.5	<b>30.4</b>
	35	8.2	<b>14.3</b>	10.5	<b>15.2</b>	20.4	<b>22.0</b>
	45	3.4	<b>6.1</b>	3.5	<b>7.0</b>	12.2	<b>14.9</b>
Ortalama	12.8	<b>16.3</b>	20.9	<b>24.9</b>	20.3	<b>21.4</b>	

TABLO II: Orsay veri kümesinde nesne tanıma test sonuçları

Nesne Tanınma Oranı	EYK (%)	10-EYK (%)
Tüm Betimleyiciler	79.09	<b>88.18</b>
En Başarılı Betimleyiciler (En iyi %50)	80.55	<b>86.36</b>

arasında eşleşmeler Hamming uzaklığına göre en yakın komşu (EYK) bulunarak belirlenmiştir. Diğer tanıma oranı ölçümünde ise en yakın on komşu (10-EYK) belirlenip, bu listeki aday referans betimleyicilerin olasılığa dayanan ve her bir referans anahtar nokta için özel olarak hesaplanan bir skorla sıralanmasıyla elde edilmiştir. Sonuçlar Tablo I'de verilmiştir.

*Dog* veri kümesi için özellikle uç açı değerlerinde tanıma oranında önemli bir artış elde edilmiştir. Örneğin, her iki yönde de 35 derece döndürülmüş görüntülerde tanıma oranı yaklaşık %75 oranında artarken 45 derecelik döndürmelerde tanıma oranı iki-üç kat artmıştır. Tüm test veri kümesi üzerinde ise ortalama tanıma oranı yaklaşık %25 oranında yükselmiştir. *Horse* veri kümesinde ise her iki yönde de 25, 35, ve 45 derecelik döndürmelerde tanıma oranında önemli bir artış elde edilip ortalama tanıma oranı yaklaşık %20 oranında yükselmiştir. *Oil* veri kümesinde ise daha az bir artış gözlenirse de tüm test veri kümesi göz önüne alındığında kullanılan yöntemin daha iyi eşleme yaptığı gözlenmiştir.

### B. Nesne Tanıma için Anahtar Nokta Eşleştirme

Betimleyici eşlemesi sadece en yakın komşu ile ve ilk on en yakın komşu ile her bir Orsay test imgesinde çalıştırılmış ve seçilen nesnelerin doğruluğu test edilmiştir. Sonuçlar Tablo II ile verilmiştir. Sadece en yakın komşu kullanılarak 110 test imgesinin 87 tanesi doğru tanınırken en yakın on komşu ile 97 test imgesi için başarılı tanıma yapılmıştır.

Betimleyiciler için istatistiksel verilerin tutulması başarıyı artırırken ek hafıza kullanımını da zorunlu kılmaktadır. Bu nedenle nesne tanıma fazlaca fayda sağlamayan betimleyiciler için istatistiklerin tutulması gereksizdir. Nesne tanıma başarımına katkı eğitim aşamında her bir betimleyici için eğitim imgeleri üzerinden hesaplanabilir. Böylelikle betimleyiciler nesne tanıma olası katkılarına göre sıralanarak en düşük katkıya sahip olanlar elenebilir. Yapılan testlerde nesne tanıma başarısında %2 kayıpla %50'ye kadar veri boyutunda azalma sağlamanın mümkün olduğunu göstermiştir. Sonuçlar Tablo II'nin ikinci sırasında yer almaktadır.

## V. SONUÇ

Çalışmada [11] tarafından düzlemsel nesnelere için önerilen gürbüz anahtar nokta eşleme yöntemi, [15] tarafından elde edilen veri seti ve yöntemler ile üç boyutlu nesnelere uyarlanmış ve eğitim verisi elde edilmiştir. Sonuçlar yöntemin üç boyutlu nesnelere için de gürbüz eşleme sağladığını göstermektedir. Ayrıca yöntem bir müze eser tanıma uygulamasına adapte edilerek farklı koşullarda nesne tanıma başarısı da ölçülmüştür.

## KAYNAKLAR

- [1] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1/2, pp. 43–72, 2005.
- [2] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, Jan 2010.
- [3] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2004.
- [4] M. Calonder, V. Lepetit, M. Özuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a Local Binary Descriptor Very Fast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1281–1298, 2012.
- [5] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: an efficient alternative to sift or surf," in *International Conference on Computer Vision*, 2011, pp. 2564–2571.
- [6] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *International Conference on Computer Vision*, 2011, pp. 2548–2555.
- [7] A. Alahi, R. Ortiz, and P. Vandergheynst, "Freak: Fast retina keypoint," in *Conference on Computer Vision and Pattern Recognition*, 2012, pp. 510–517.
- [8] T. Trzcinski, M. Christoudias, and V. Lepetit, "Learning image descriptors with boosting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 597–610, March 2015.
- [9] G. Levi and T. Hassner, "LATCh: learned arrangements of three patch codes," *CoRR*, vol. abs/1501.03719, 2015. [Online]. Available: <http://www.openai.ac.il/home/hassner/projects/LATCh>
- [10] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 20, no. 2, pp. 91–110, 2004.
- [11] F. Uzyıldırım and M. Özuysal, "Instance detection by keypoint matching beyond the nearest neighbor," *Signal, Image and Video Processing*, vol. 10, no. 8, pp. 1527–1534, 2016.
- [12] V. Balntas, L. Tang, and K. Mikolajczyk, "Bold-binary online learned descriptor for efficient image matching," in *Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2367–2375.
- [13] F. Li and J. Kosecka, "Probabilistic location recognition using reduced feature set," in *International Conference on Robotics and Automation*, 2006.
- [14] M. Dymczyk, S. Lynen, M. Bosse, and R. Siegwart, "Keep it brief: Scalable creation of compressed localization maps," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2015, pp. 2536–2542.
- [15] P. Moreels and P. Perona, "Evaluation of features detectors and descriptors based on 3D objects," *International Journal of Computer Vision*, 2006.
- [16] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [17] A. Andoni and P. Indyk, "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," *Communications of the ACM*, vol. 51, no. 1, pp. 117–122, 2008.
- [18] O. Chum and J. Matas, "Matching with prosac - progressive sample consensus," in *Conference on Computer Vision and Pattern Recognition*, San Diego, CA, June 2005, pp. 220–226.