

Öğrenme Tabanlı Otomatik Teleoperasyon Sistemi

Learning Based Automated Teleoperation System

Hanife Coşkun, Şeyma Yücer, Abdullah Akay ve Yusuf Sinan Akgül

GTU Vision Lab, <http://vision.gyte.edu.tr>
Bilgisayar Mühendisliği Bölümü
Gebze Teknik Üniversitesi,
Kocaeli, 41400, Türkiye
{hcoskun, syucer, aakay, akgul}@gtu.edu.tr

Özetçe—Teleoperasyon sistemlerinin, operatörden kaynaklanan hatalar ve sistemdeki gecikmelerden dolayı verimsiz olduğu bilinmektedir. Bunun yanında sürekli tekrarlanan iş parçacıkları barındıran teleoperasyon görevleri verimsizliği daha da arttırmaktadır. Bu çalışmada söz konusu problemleri ortadan kaldırmak için operatör tarafından gerçekleştirilen hareketlerin ve sahnedeki nesnelerin otomatik öğrenilmesi önerilmiştir. Öğrenilen hareket ve nesne bilgilerine göre, sistemin anlık durumuna en uygun kararların alınıp uygulanması amaçlanmıştır. Geliştirilen sistemi sınamak için bir test ortamı oluşturulmuş ve bir dizi deney gerçekleştirilmiştir. Yapılan deneylerde geliştirilen sistemin teleoperasyon sistemlerinin bahsedilen sorunlarına kısmen çözüm getirmesinin yanında operatör performansını da arttırdığı gözlemlenmiştir.

Anahtar Kelimeler — Teleoperasyon, Denetimsiz Öğrenme, Makine Öğrenmesi, Bilgisayarla Görme, RGB-D kameralar, Kinect, Üç Boyutlu (3B) Geri Çatım

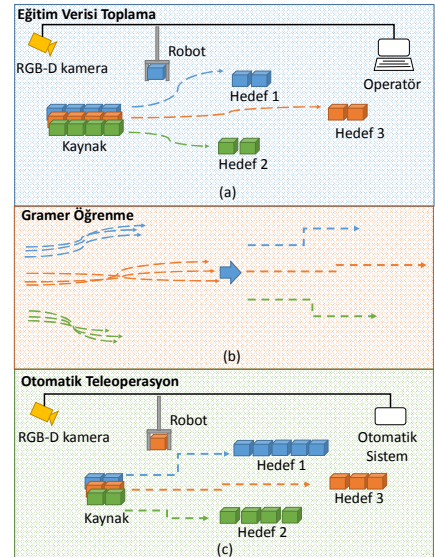
Abstract—Teleoperation systems are inefficient due to operator errors and system delays. Moreover, complexity is even higher for teleoperation tasks that involve repetitions. In this paper a technique that consists of automatic learning of operator commands and scene objects is proposed. Exploiting learned action patterns and scene objects, system is able to take the appropriate actions for its current state. A number of experiments prepared to assess the proposed system. Experimental results show that the system can solve some known problems of teleoperation and increase the throughput of the operators.

Keywords— Teleoperation, Unsupervised Learning, Machine learning, Computer Vision, RGB-D cameras, Kinect, 3D Reconstruction

I. GİRİŞ

Teleoperasyon sistemlerinde operatörler uzaktaki araçları kontrol ederek belirli görevleri yerine getirir [1, 2]. Daha özel bir durum için işlek bir limanda ticaret gemilerine konteyner yüklemesi yapan bir vinç operatörünü düşünelim. Uzman operatör liman içerisindeki birden fazla sayıda bulunan vinçlerin hepsini kontrol etmelidir. Operatörün daha verimli çalışabilmesi için bir merkezden bütün vinçleri yönetmesi gerekmektedir. Dolayısıyla uzaktan kontrol ile

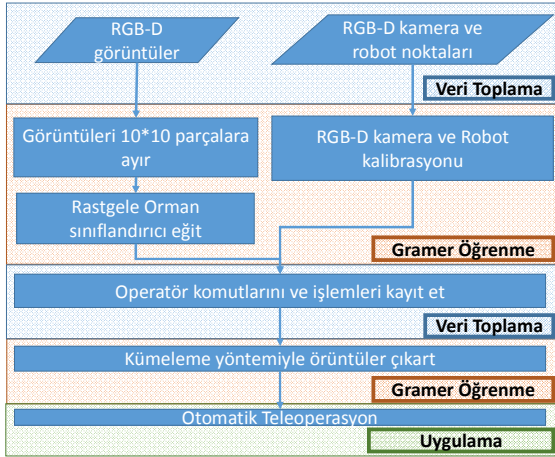
ilgili birçok teleoperasyon güçlüğü söz konusudur. Üstelik teleoperasyondaki bu zorluklar, kısa süreli işler için büyük sorun teşkil etmese de sürekli tekrarlanan iş parçacıklarının içeren uygulamalarda büyük bir zaman kaybına sebep olmaktadır [3]. Operatörün bir gemiye yükleme işlemini tamamlamak için yüzlerce konteyneri vinç vasıtasıyla taşıması icap eder. Her bir konteyner taşınırken teleoperasyon güçlükleri sebebiyle zaman kaybedilir ve bütün taşıma işi tamamlandığında görev oldukça verimsiz bir şekilde yerine getirilmiş olur [4].



Şekil 1. Geliştirilen sistemin genel yapısı. (a) Operatörün kontrol ettiği robottan veri toplanması. (b) Toplanan verilerden hareket kümelerinin çıkarılması. (c) Öğrenilen hareketlerle teleoperasyon görevinin otomatik olarak tamamlanması.

Bu çalışmada, geliştirilen sistem sayesinde sürekli benzer hareketlerin yapıldığı teleoperasyon uygulamalarında operatörün hareketlerinden otomatik olarak neleri nasıl yaptığının öğrenilmesi ve kalan görevlerin yarı otonom yapılması amaçlanmıştır. Böylelikle operatör robotu bir süre kullanarak ona yapması gereken hareketleri öğretecektir ve geri kalan zamanını diğer görevleri tamamlamak için kullanabilecektir.

Önerilen sistemin genel yapısı Şekil 1’de görülmektedir. Buna göre, sistem 3 fazdan oluşmaktadır; veri toplama, gramer öğrenme ve uygulama (otomatik teleoperasyon) fazları. Veri toplama fazında operatör robotu kullanarak teleoperasyon görevini sürdürür. Bu süre zarfında sistem, robot ve sahne verilerini toplar. Toplanan verilerden gramer öğrenme fazında makine öğrenmesi yöntemleriyle üst seviye bilgiler çıkartılır ve sistemin işleyiş dinamikleri yani operatörün hangi iş parçacıklarını nasıl yaptığı bilgisi elde edilir. Ardından uygulama fazında robot, sistem tarafından öğrenilen işleyiş dinamikleri sayesinde tamamen otonom kontrol edilir ve bu şekilde teleoperasyon görevi tamamlanır.



Şekil 2. Sistemin genel akış diyagramı

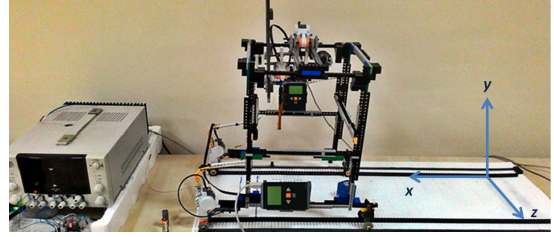
Şekil 2’de geliştirilen sistemin akış diyagramı verilmiştir. Buna göre sistemin sahnedeki nesnelere otomatik tanınması için gerekli eğitim verisi RGB-D (renk ve derinlik) görüntülerinden elde edilir. Bunun yanında robot referans dünyası ile RGB-D kameranın referans dünyasını ilişkilendirmek için bir dizi nokta eşleşmesi elde edilir. Sonrasında her bir RGB-D görüntüsü 10x10 parçalara ayrılır ve parçalar kullanılarak Rastgele Orman tabanlı bir sınıflandırıcı eğitilir. Robot ve RGB-D kamera referans dünyasını ilişkilendirmek için ise 3B dünyalar arası dönüşüm bulan bir yöntem kullanılır. Ardından operatörün robotu elle kontrol etmesi esnasında verilen komutlar ve işleme tabi tutulan nesnelere kayıt edilir. Operatörün komutlarından bir kümeleme yöntemi ile örüntüler çıkarılır. Son olarak RGB-D kamera ile analiz edilen sahne üzerinde çıkartılan örüntüler kullanılarak robotun otomatik yönetimi gerçekleştirilir.

Bu bildirinin artı kalan kısmında öncelikle basitleştirilmiş robot düzeneği anlatılacak, sonrasında gramer öğrenme ve sahne analizi modülleri açıklanacaktır. Ardından deney sonuçları sunulacak ve çalışmanın çıktıları tartışılacaktır.

II. BASİTLEŞTİRİLMİŞ ROBOT DÜZENEGİ

Bu çalışmada, önerilen sistemi gerçek boyutlarda inşa etmek pratik olmayacağından sistem küçük ölçekli bir robot kullanılarak geliştirilmiş ve test edilmiştir. Lego Mindstorms parçaları kullanılarak basitleştirilmiş bir robot düzeneği geliştirilmiştir (Şekil 3). Düzenekte 4 porttan

dinlenebilen 4 servo motor, 2 NXT cihazı, bir güç kaynağı ve Arduino Uno ile kontrol edilen elektromıknatis bulunmaktadır. Robotun 3 boyutta hareketini sağlamak amacıyla x eksenine 2 motor y ve z eksenlerine 1’er motor yerleştirilmiştir. Sistemin ilk durumundan son durumuna kadar hareketinin kaymamasını ve devamlılığını sağlamak için platform sınırları belirlenmiştir.



Şekil 3. Basitleştirilmiş robot düzeneği

III. DENETİCİSİZ ÖĞRENME BİRİMİ

Robotun platform üzerinde konumlanmış nesneyi alıp başka bir yere bırakma işlemi, hareket örneklerini; bu işlemin defalarca yapılması ise işlenmemiş ilk veri kümelerini oluşturmaktadır. Alınan ham veri kümesi hangi yönde ne kadar hareket ettiğinin sıralanıp numaralandırılmasıyla oluşturulmuştur. Oluşan bu verinin histogram değerleri; toplam verinin histogramı, ilk çeyrek histogramı, ikinci çeyrek histogramı, üçüncü çeyrek histogramı ve sonuncu çeyrek histogramları olarak ayrı ayrı hesaplanıp art arda yerleştirilmiştir. Bu şekilde, yapılan hareketlerin sırası ve sayısı dikkate alınarak öğrenme işlemine hazır veri kümesi oluşturulmuştur.

K-Merkezli [5] bilinen kümeleme sorununu çözmek için kullanılan denetimsiz öğrenme algoritmalarından biridir. Robot hareketlerinin kümelenmesinde bu teknik kullanılmıştır. Ana fikir k adet noktayı, küme içinde alt küme merkezi olarak tanımlamaktır. K-merkezli kümeleme, veri içindeki kümelerin bulunması içindir ve kümeler, asıl örneklerine karşılık gelen merkezleri tarafından gösterilir.

K-Merkezli kümeleme yöntemi küme sayısını girdi olarak almaktadır. Fakat bu problemde küme sayısı dinamik olarak değiştiği için böyle bir bilgi sağlanamamaktadır. Dolayısıyla küme sayısını otomatik olarak bulan bir yöntem kullanılmalıdır. Bunun için 1 den n ‘e kadar k değerleri için tüm hareket verilerinin yoğunluk değerleri bulunmuştur.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

Denklem (1)’de yoğunluk değerinin tanımı yapılmıştır [6]. Doğru kümeleme için yoğunluk değerleri sıralanmıştır. Yoğunluk değer değişiminin minimuma indiği yerdeki k değeri doğru küme sayısını vermektedir.

Ölçümleme(kalibrasyon) tanımlanmış koşullar altında izlenebilirliği sağlanmış sistemlerin arasındaki ilişkiyi belirleyen işlemler dizisidir. İki ayrı dünya olarak tanımlanan sistemlerin etkileşimleri için ölçümleme işlemi yapılır. Kinect’in aldığı görüntünün piksel değerlerinden, gerçek hayattaki değerleri arasındaki ilişkinin bulunması önemlidir. Çünkü sistemin Kinect’ten görüntü olarak

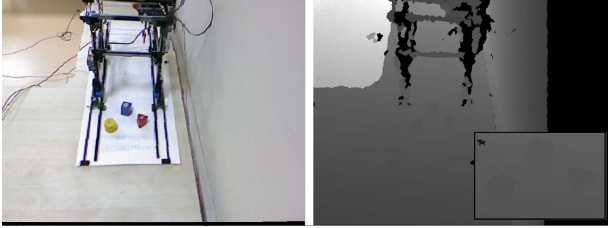
hareket ettirilecek cisimlerin konumlarını tespit etmesi gerekmektedir. Bu da bir noktanın iki farklı dünyadaki değerlerinin arasında ki yer değiştirme ve dönüşüm matrislerinin bulunması ile çözülebilir [7]. Robot düzeneğinin kendi içinde bulunduğu koordinat sisteminden ve Kinect dünyasından örnek nokta verileri alınarak ilişkili matrisler hesaplanır. Sonrasında gelen her yeni nokta verisi için bu matrisler kullanılarak iki dünya arasında ki etkileşim sağlanır.

IV. SAHNE ANALİZİ VE ETKİLEŞİMİ

Bir RGB-D kamerası vasıtasıyla etkileşime girilerek 3B gerçek dünya ortamı hakkında bilgi elde edilmeye çalışılmıştır. Bu bilgi sahnedeki etkileşim nesnelere verilerini kapsamaktadır. Elde edilen bu bilgiler sahneyi robot kolu aracılığıyla manipüle etmek için kullanılacaktır.

Sahnedeki nesnelere 3B gerçek dünya ortamı hakkında bilgilerini elde etmek amacıyla Kinect ile sahnenin RGB-D görüntüsü elde edilir [8].

Kinect üzerindeki renk ve derinlik kameralarının arasında belli bir mesafe bulunduğundan, iki kamera ile elde edilen görüntüler üst üste konulduğunda bire bir örtüşmeyeceklerdir. Dolayısıyla renkli görüntü üzerindeki (x,y) pikselinin derinliği öğrenilmek istendiğinde derinlik görüntüsündeki (x,y) pikselinden farklı bir noktanın derinliği elde edilecektir. Bu problemi çözmek için Kinect'ten alınan renk ve derinlik görüntülerinin örtüşürülmesi gerekmektedir (registration) [9]. Örtüşürülme işleminden sonra girdi görüntüleri Şekil 4'deki gibi olacaktır.



Şekil 4. Örtüşürülmüş renk ve derinlik görüntüsü

Sahnede kırmızı üçgen prizma, mavi dikdörtgenler prizma ve sarı silindir olmak üzere 3 farklı tip nesne mevcuttur. Bu nesnelere tanımak için örtüşürülmüş RGB görüntüsüne Rastgele Orman sınıflandırıcısı uygulanarak görüntü üzerindeki nesnelere yerleri tespit edilir.

Rastgele Orman, sınıflandırma için kullanılan topluluk algoritmalarından biridir. Bu yöntem sınıflandırmanın doğruluğunu artırmak amacıyla birden fazla karar ağacından oluşur ve her bir karar ağacı bir sınıf etiketi için oy verir. Bir örnek, karar ağacı tahminlerinde en çok oyu alan sınıfa atanarak sınıflandırılır [10].

Sahnede olabilecek olan 3 nesnenin 5'er derecelik açılar ile 12 farklı uzaklıktan alınan görüntüleri ve arka plan görüntüsü, 10x10 pencere boyutunda küçük görüntülere parçalanarak eğitim kümesi oluşturulur. Eğitim kümesi ile Rastgele Orman yöntemindeki karar ağaçları eğitilir.

Nesnelere tanıma işleminde ise Kinect ile alınan görüntüler sonucu elde edilen örtüşürülmüş RGB görüntüsü üzerinde 10x10 boyutlu pencereler halinde gezilerek her pencere için Rastgele Orman sınıflandırıcı uygulanır. Bu işlem sonucunda 10x10 boyutundaki pencerenin üçgen prizma, dikdörtgenler prizma, silindir ya da arka plan seçeneklerinden biri olan bir sınıf etiketi tahmin edilir. 10x10 boyutlu pencerenin orta noktasındaki piksel tahmin edilen sınıf için belirlenen değer atanarak bir etiket görüntüsü elde edilir. Bu görüntü üzerinde 15x15 boyutunda pencereye sahip medyan filtrelemesi uygulanarak gürültü azaltılır. Etiket görüntüsü üzerinde bütünleşik olarak bulunan aynı sınıf etiketlerine ait alanlar bulunarak nesnelere görüntü üzerindeki yaklaşık yerleri tespit edilir.

Buraya kadar nesne tanıma işlemi için sadece 2B renk bilgisi kullanılmış ve görüntü üzerindeki nesnenin 2B bilgisi elde edilmiştir. Fakat elde edilen bu sonuç ilgili nesnenin 3B pozisyonun bilgisini vermemektedir. Bu bağlamda bölütlenen 2B nesnenin piksel ve derinlik bilgileri kullanılarak parçalı 3B nokta bulutuna dönüştürülmüştür (Şekil 5).



Şekil 5. Nesnenin 3B nokta bulutu

Elde edilen parçalı 3B nokta bulutuna en iyi oturan dikdörtgenler prizma masraf algoritması kullanılarak bulunmuştur. Bulunan prizmanın merkez noktası 3B nokta bulutunu temsil eden nesnenin, gerçek dünya koordinatı olarak kabul edilmesiyle geri çatım yapılmıştır.

V. DENEYLER

Otomatik teleoperasyon için tasarladığımız düzenek üzerinde beş adet deney gerçekleştirilmiştir. Deneylerden ilki, hareketlerin öğrenilmesinde kullanılan k-merkezli kümeleme algoritmasının doğru kümelemesi için yoğunluk değerlerinin denenmesidir. Amaç, beklenen k değerlerini en yüksek başarıyla veren yoğunluk sınırının bulunmasıdır.

Tablo 1. Yoğunluk-doğru kümeleme olasılığı ilişkisi

Yoğunluk Üst Limiti	Doğru Kümeleme Oranı	Ortalama Hata
0.016	%55.71	14.062
0.008	%72.29	6.172
0.004	%81.86	2.758
0.002	%90	1.351
0.001	%100	0.616
0.0005	%100	0.616

Tablo 1'de kümelerin merkezleri için yoğunluk değeri arttıkça doğru kümeleme olasılığının düştüğü görülmektedir. Küme sayıları yüksek olan veri kümelerinde hassasiyet azaldıkça kümeleme hataları artmaktadır.

Deneylerden ikincisinde seçilen üç kümeleme algoritmasının başarılarının karşılaştırılması amaçlanmıştır.

Tablo 2’de farklı küme sayılarındaki veri kümeleri için test sonuçları gösterilmiştir. Kullanılan üç algoritma için veri kümesindeki hareketlerin merkezleri ile hareketlerin kendileri arasındaki Manhattan uzaklık değerleri hesaplanmıştır. Ortalama uzaklık hatasının küme sayısı arttıkça arttığı, ancak bu artışın en az k-merkezli kümelemede olduğu gözlemlenmiştir.

Tablo 2. Deneticisiz kümeleme algoritma hataları

Küme sayısı	3	5	8	10	15
K-Merkezli Kümeleme	0.45	0.30	0.28	1.08	0.74
En Uzakla Kümeleme	0.97	0.98	0.49	3.39	2.82
Beklenti Maksimizasyonu	1.63	0.81	0.35	2.31	9.80

Üçüncü deneyde görüntü üzerindeki nesnelerin tespiti için kullanılan Rastgele Orman sınıflandırıcısının başarısı test edilmiştir. Bu sınıflandırma sistemini test etmek için; hem birbirine benzeyen hem de değişik geometrik şekil ve büyüklükte olan 10 adet nesnenin görüntülerinden bir veri kümesi oluşturulmuştur (Şekil 6).



Şekil 6. Sınıflandırma testi için kullanılan nesneler

Rastgele Orman sınıflandırıcısı, 10-katlı çapraz doğrulama yöntemi ile test edilmiş ve elde edilen sonuçlar Tablo 3’de gösterilmiştir.

Tablo 3. Rastgele Orman yönteminden elde edilen sonuçlar

Test Kümesi Ort. Örnek Sayısı	Doğru Sınıflandırılmış Ort. Örnek Sayısı	Ort. Başarı Yüzdesi(%)	Standart Sapma
282.2	259.8	91.73	1.62

Dördüncü deneyde nesnelerin parçalı nokta bulutundan 3B pozisyonlarını bulma yöntemi test edilmiştir. Kinect ile alınan 50 görüntü üzerindeki nesnelerin nokta bulutundan hesaplanan 3B pozisyonları, kesin referans pozisyonları ile karşılaştırılmış ve ortalama hata 9.04 mm, standart sapma ise 10.36 mm olarak hesaplanmıştır. Hesaplanan ve kesin referans pozisyonları arasındaki farkların kaynağının, Kinect’ten alınan görüntünün düşük çözünürlükte olması ve Kinect’in kendi ölçümlemesindeki hatalar olduğu değerlendirilmektedir.

Son deneyde ise sistem üzerinde elle yapılan hareketlere kıyasla önerilen yöntem ile otomatik öğrenilmiş hareketlerin

gerçekleştirilmesinin daha az zaman aldığı gösterilmek istenmiştir. Bunun için üç farklı kümeye ait 12 adet hareketi içeren temsili bir görev, klavye kullanılarak sistem üzerinde gerçekleştirilmiş ve bu iş için harcanan süre 757 sn. olarak hesaplanmıştır. Sonrasında, bu 12 hareketin ilk 6 hareketi (her kümeden iki adet), klavye ile yapılarak sisteme öğretilmiştir. Kalan 6 hareket, öğrenilen hareket kümeleri kullanılarak sistem tarafından otomatik olarak yapılmıştır. Otomatik teleoperasyon ile bu görev 567 sn. içerisinde tamamlanmıştır. Gerçekleştirilen bu testte, otomatik teleoperasyon sayesinde %25.09 oranında zaman kazancı sağlandığı gösterilmiştir.

VI. SONUÇ

Bu çalışmada sürekli tekrarlanan ve birbirine benzer iş parçacıkları içeren teleoperasyon görevlerinin verimsizlik problemlerine çözüm getiren makine öğrenmesi tabanlı bir yöntem geliştirilmiştir. Önerilen sistemin, oluşturulan deney ortamında bir dizi test ile sınanarak başarılı sonuçlar ortaya koyduğu görülmüştür. Bu deneylerde sistemin teleoperasyon görevlerini otomatik gerçekleştirebilmesi sayesinde operatörlerin verimliliğini hatırı sayılır derecede iyileştirildiği gözlemlenmiştir. Dünya ölçeğinde uluslararası taşımacılık, inşaat gibi sektörlerin hızla geliştiği göz önüne alındığında önerilen sistemin gelecekte geniş kullanım alanı bulacağı düşünülmektedir. Gelecekte yapılması planlanan çalışmalar arasında sistemin sahne hakkında hiçbir bilgi sahibi olmaksızın eğitilmesi ve otomatize edilmesi vardır.

KAYNAKÇA

- [1] Sato T., Hirai S., "Language-aided robotic teleoperation system (LARTS) for advanced teleoperation," in *Robotics and Automation*, 1987.
- [2] A. Kheddar, C. Tzafestas, P. Coiffet., "The hidden robot concept-high level abstraction teleoperation," in *International Conference on Intelligent Robots and Systems*, 1997.
- [3] Brenna D. Argall, Sonia Chernova, Manuela Veloso ,Brett Browning, "A Survey of Robot Learning from Demonstration," in *Robotics and Autonomous Systems*, 2009.
- [4] Andrew Y. Ng, H. Jin Kim, Michael I. Jordan, Shankar Sastr, "Inverted autonomous helicopter flight via reinforcement learning," in *International Symposium on Experimental Robotics*, 2004.
- [5] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability," in *University of California Press*, Berkeley, 1967.
- [6] D. T. Pham, S. S. Dimov and C. D. Nguyen, "Selection of K in K-means Clustering," in *Cardiff University*, Cardiff, 2004.
- [7] N. HO, "Finding Optimal Rotation and Translation Between Corresponding 3D Points".
- [8] Smisek, Jan, M. Jancosek and T. Pajdla, "3D with Kinect," in *Consumer Depth Cameras for Computer Vision*, London, 2013.
- [9] P. Henry, M. Krainin, E. Herbst, X. Ren and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," in *The International Journal of Robotics Research*, 2012.
- [10] L. Breiman, "Random Forests," in *Machine Learning*, 2001.