Etkileşimli Görsellik Kullanılarak El Aletlerinin Sağlamlık Tahmini Affordance Prediction of Hand Tools Using Interactive Perception

Yiğit Çalışkan^{1,2}, Pınar Duygulu¹, Erol Şahin²

¹Bilgisayar Mühendisliği Bölümü
Bilkent Üniversitesi, Ankara, Turkey

caliskan@cs.bilkent.edu.tr, duyugu@cs.bilkent.edu.tr

²KOVAN Araştırma Laboratuvarı, Bilgisayar Mühendisliği Bölümü
Orta Doğu Teknik Üniversitesi, Ankara, Turkey

erol@ceng.metu.edu.tr

ÖZETÇE

Günlük hayatta bir iş için seçilecek el aleti genellikle o aletin dış görünüşü ve nesneler üzerindeki etkisinden yola çıkılarak seçilir. Bu etki, seçilen aletin sağlamlığı (ing. affordance) belirler. Bu çalışmanın amacı, el aletlerinin sadece dış görünüşlerini kullanarak onların sağlamlıklarını belirlemek ve insansı robotların basit alet kullanımı için bir zemin oluşturmaktır. Bu amaçla yapılan çalışmada insan tarafından etkileşilen el aletlerinin fonksiyonel bölgelerinin düzlük, sivrilik, iki uç arası uzaklık, gri ölçek renk histogramı gibi öznel özellikler çıkartılarak belli sağlamlık modelleri eğitilmiştir. Sağlamlıkları öğrenilmek istenen el aletinden çıkarılan her bir öznel özellik, eğitilmiş modellere verilerek bu aletin kesebilme, delebilmek, sıkıştırılabilir ve itirebilir sağlamlıklarından hangilerine sahip olduğu belirlenir. Testlerde modelin %93.1 oranında sağlamlıkları doğru tahmin ettiği görülmüştür. Bu sonuçlara bakılarak modelin insansı robotlarda basit alet kullanımı için uygun bir zemin oluşturduğu söylenebilir.

ABSTRACT

In daily life, the selection of a hand tool for a job depends on appearance of the tool and its effect on the objects. The effect determines the affordance of the chosen tool. Aim of this work is to determine the affordances of hand tools based only on their appearance and to build a basis for simple tool usage of humanoid robots. Towards this end, in this work from the functional regions of human interacted hand tools, sharpness, bluntness, distance between two tip and grayscale histogram features are extracted and specific affordance models are trained. The features of a hand tool which its affordances wanted to be learned are given to the trained models to determine which affordances that the tool has like can cut, can push, can squeeze, can pierce. During testing, the model predicted the affordances %93.1. From this results it can be said that, this model sets a basis for simple tool usage of humanoid robots.

1. GİRİŞ

İnsanlar günlük yaşamlarında işlerini daha da kolaylaştırmak amacıyla el aletlerini kullanırlar. Yapılacak işe uygun aleti seçmek işin başarıyla gerçekleşmesinde büyük rol oynar. Her aletin bir veya birden fazla sağlamlığı (ing. affordance) vardır. Sağlamlık terimi ilk olarak J.J. Gibson tarafından ortaya atılmıştır. Gibson, sağlamlığı bir nesnenin bir organizmaya herhangi bir ortamda sunduğu eylemler olarak tanımlamaktadır. Gibson'a göre sağlamlık özne tarafından doğrudan gözlenebilen bir kavramdır, bir diğer deyişle ortamda bulunan nesnelerin anlamlarının doğrudan anlaşılmasıdır[1]. Mesela yatay ve katı bir yüzeyin yürünebilir olmasının, belli bir ağırlıkta olan taşın fırlatılabilir olmasının anlaşılması gibi[2]. Bu çalışmada ise el aletlerinin sağlamlıklarının anlaşılması hedeflenmektedir. Mesela, bir makasın bir kağıdı ikiye bölmesi sonucundan, makasın kesebilirlik özelliğinin olduğunun öğrenilmesi gibi. Bu sağlamlıklar aletin fiziksel özelliklerine ve aletin uygulanacağı hedef nesneye bağlıdır. Amacımız, belirli aletlerle, denemeyle öğrenmek de denilebilen etkileşimli görsellik kullanarak aletlerin fiziksel özelliklerini öğrenmek ve bu özelliklerle el aletlerinin ne gibi sağlamlıklara sahip olduğunu hedef nesneler olmadan öğrenmek ve Şekil 1'deki insansı robot iCub'da bu çalışmadaki karmaşık aletler olmasada daha basit aletlerin etkileşimli kullanımının temeli oluşturmaktır.

Şekil 1: İnsansı Robot iCub

2. İLGİLİ ÇALIŞMALAR

Etkileşimli görsellik ve aletlerin sağlamlığı ile ilgili çalışmalar genellikle alet ve aletin uygulanacağı hedef nesne üzerine kuru-ludur. Shinci ve diğerleri çalışmalarında nesne kavramını, nesneler ve onların uygulanacağı hedefler arasındaki etkileşimden doğan sonuçları kullanarak oluşturmuşlardır[3]. Bu çalışmada, aletlerin şekillerini ve hedefin eylem öncesi ve sonrasındaki değişikliklerini öznitelik olarak kullanmışlardır. Nakamura ve diğerleri bu çalışmayı daha da ilerletip nesneden ve hedeflerden çıkarılan özniteliklerin yanısıra aletlerin kullanımları sırasındaki el ve tutuş şekillerini de işin içine katarak aletlerin tutulan bölümlerini tespit etmişlerdir[4]. Katz ve diğerleri ve Willimon ve diğerleri ile çalışmalarında hedef nesne kullanmadan, nesneler üzerinde etkileşimli görsellik kullanarak nesnelerin yapıları hakkında eklem bilgisi, nesne iskeleti gibi nitelikleri elde etmişlerdir[5][6][7]. İleriki bölümlerde açıklanacak olan sunulan modelin özelliği ise hedef nesne kullanılmadan aletler üzerinde etkileşimde bulunup sadece görsel bilgileri kullanarak aletlerin belirli kısımlarının sağlamlıklarını bulmaktır.

3. MODEL DETAYLARI

Sunulan model, girdi olarak öznenin alet üzerinde etkileşiminin görüldüğü bir video dosyasını (Şekil 2) alıp işleyerek, çıktı olarak etkileşimde bulunulan aletin sağlamlık tahminlerini verir. Sistem, yöntemlerin uygulanmasında kolaylık sağlaması amacıyla birkaç varsayım üzerine kurulmuştur. Bu varsayımlar:

Şekil 2: Video Dosyasını Oluşturan Resimlerden Örnekler

El aletleri siyah yüzey üzerinde paralel, fonksiyonel kısmı sola bakacak ve ağzı kapalı şekilde konumlandırılmalıdır. El aletleri en fazla bir eklemle sahip olmalıdır. Son olarak, özniteliklerin kolayca izlenmesi amacıyla aletlerin üzerinde renkli bantlar kullanılmıştır. Konum varsayımı fonksiyonel bölgenin bulunmasında kolaylık sağlamaktadır, ancak aletlerin tutulma bölgelerinin tespiti gibi ekstra bir yöntem bu varsayımı kaldırmada yardımcı olabilir. Modelin genel işleyişi kabaca Şekil 3'de ve modelin kısımları Şekil 4'deki makas örneği üzerinde görülebilir.

3.1. Aletlerin Eklem Tespiti

Modelin ilk kısmı, etkileşimde bulunulan aletin bir eklemle sahip olup olmadığının bulunmasıdır. Eklem tespiti edilmesi, ilgili fonksiyonel bölgelerin çıkartılmasında kolaylık sağlar. Eklem tespiti için optik akış (ing. optical flow)[8] metodu kullanılmıştır. Optik akış metodunda kullanılacak olan özniteliklerin kolaylıkla bulunabilmesi ve izlenebilmesi amacıyla el

aletlerinin üzerlerine renkli bantlar yapıştırılmıştır. Eklem tespit algoritması, aletle etkileşim esnasında eklemdeki özniteliklerin, eklem dışındaki özniteliklere olan uzaklığının hareket boyunca değişmemesi gerçeğini kullanarak çalışmaktadır. Eklem koordinatları hesaplamalardaki parazitler göz önüne alınarak, değişimin en düşük olduğu dört öznitelik pozisyon ortalaması alınarak bulunmaktadır. Eğer ki tüm öznitelikler arasında oldukça gözdürülebilecek bir uzaklık değişimi olursa bu, aletin eklemle sahip olmadığını gösterir.

3.2. Aletlerin İlgili Bölgelerinin Çıkarımı

Eklem tespitinin sonucunda bir eklem bulunduysa, aletin konumu, aletin ağzının maksimum ve minimum genişliği göz önünde bulundurularak ilgili bölgeler belirlenir. İlgili fonksiyonel bölge örnekleri Şekil 4'teki örnekte görülebilir. Eğer bir eklem bulunamadıysa, ilgili bölge, aletin sol yarısı olarak belirlenir. Bu belirleme işlemlerinde ilgili bölgenin boyutu, aletin siyah-beyaz resim formatları kullanılarak belirlenmiştir. İlgili bölgelerin belirlenmesi işleminde aletlerin sola bakacak şekilde konumlandırılma varsayımı kullanılmıştır.

3.3. Aletlerin Sınırlarının Belirlenmesi

İlgili bölgelerin belirlenmesi, resimlerdeki gereksiz ayrıntılardan kurtulmayı sağlamakta ve gerekli öznitelik çıkarım işlemlerini doğrudan uygulamamıza imkan vermektedir. Öznitelik çıkarım işlemleri için ilgili bölgelerin gri ölçekli resimleri kullanılmaktadır. Bu resimler gri ölçek yoğunluk histogramlarının karşılaştırması, uç bölgelerin belirlenmesi ve uç bölgelerin düzlük ve sivrilik gibi özniteliklerin bulunmasında kullanılmaktadır. Ancak gri ölçekli resimlerde öznitelikler çıkarılmadan önce aletlerin dış sınırlarının belirlenmesi gereklidir. Bu amaçla OpenCV[9] kütüphanesinin dış sınır çıkartma algoritmaları kullanılmıştır. Ancak tahmin edildiği üzere sınırları bulmak için kullanılan eşik değerleri her alette aynı sonucu vermemekte bu nedenle sınırlarda bazı boşluklar kalabilmektedir. Bunu engellemek amacıyla konveks kabuk (ing. convex hull)[10] metodu kullanılarak dış noktaların birbirine bağlanması sağlanmış ve boşluklar kapanmıştır. Böylece resimler düzlük, sivrilik ve uç nokta özniteliklerinin çıkarılması için sınırlar belirli ve basit hale getirilmiştir.

3.4. İlgili Bölgelerde Düzlük ve Sivrilik Özniteliklerinin Belirlenmesi

Sınırın belirli olduğu siyah-beyaz resim formatı kullanılarak eğer el aleti bir eklemle sahip ise aletin ağzı kapalı ve ağzının maksimum genişlikte olduğu iki resim düzlük ve sivrilik öznitelikleri için değerlendirilir. Bu öznitelikler bir aletin herhangi bir nesneyi ittirebilme veya delebileme sağlamlıklarını tespit etmeye yardımcı olur. Eğer alet bir eklemle sahip değilse, aletin herhangi bir siyah-beyaz resim formatı öznitelik değerlendirmesi için kullanılır. Bu öznitelik çıkarım işleminde ilk olarak konveks kabuk sınırlarının iki boyutlu koordinatları sırayla elde edilir. Her iki noktada bir bu koordinatlardan örnek alınır ve eğim hesaplaması için kullanılır. Hesaplanan eğimlerin arc-tanjantları alınarak örneklenen ikişer noktalar arasındaki dereceler bulunur. Bir konveks kabuk için hesaplanan derece dizisi arasından maksimum derece bulunur ve bu dizi içindeki maksimu-

Şekil 3: Modelin Ana Kısımları

Şekil 4: Modelin Örnekli Ana Kısımları

mum dereceler toplanır ve ardından öznitelik olarak kullanılır. Eklemler bir el aletinde bu işlemin sonucunda el aletinin ağzının kapalıyken elde edilen düzlük-sivrilik özniteliği ve ağzının maksimum açık olduğu resimdeki, ağzın iki parçasından alınan öznitelik değerlerinin ortalamasının ortalaması alınarak öznitelik değeri elde edilir. Eklemsiz bir alette de işlemin sonucunda, benzer şekilde siyah-beyaz tek bir resimden elde edilen değer öznitelik olarak kullanılır.

3.5. İlgili Bölgelerde Uç Noktalar Arasındaki Uzaklığın Bulunması

Uç noktalar arası uzaklık bulma işlemi sadece eklemlerde ve maksimum ağız genişliğine sahip siyah-beyaz resim formatında kullanılır. Bu özniteliğin bulunma amacı bir aletin herhangi bir nesneyi sıkıştırıp sıkıştıramayacağını belirlemektir. Bu uzaklığın bulunması için öncelikle ağızların iki ucunun koordinatları bulunmalıdır. Bunu gerçekleştirmek için bir önceki bölümde bulunan maksimum derece dizisinin ortasındaki derece bulunur ve koordinatları uç nokta olarak kabul edilir. Bu işlem sağ ve sol olmak üzere her iki ağız için gerçekleştirilir. İşlem sonunda uzaklık, elde edilen iki koordinata uygulanan öklid uzaklığı formülü ile bulunur.

3.6. İlgili Bölgelerin Gri Ölçek Yoğunluk Histogramlarının Karşılaştırılması

Yoğunluk histogramlarının karşılaştırılması işlemi bir el aletinin kesici bir özelliğinin olup olmadığını tespit etmek için

kullanılır. Eklemler ve kesici bir özelliği olan bir aletin kesme işlemini gerçekleştirmesi için sağ ve sol ağzının birbirine sürtünerek üst üste gelmesi gereklidir. Bu nedenle kesici özelliği olan bir aletin ağzının açık ve kapalı olduğu resimlerin gri renk ölçeklerinde yoğunluk farkı olur. Bunu tespit etmek amacıyla aletin ağzının kapalı olduğu ve maksimum açıklığa sahip resimler kullanılır. Her iki resimden de gri ölçek yoğunluk histogramı elde edilir ve bu histogramlar arasındaki kesişim değerleri bulunarak bir öznitelik değeri elde edilir.

3.7. Model Eğitimi

Modelin eğitimi için işlemler sonucunda elde edilen öznitelik değerleri kullanılır. Model için bir aletin bir nesneyi keşebilirliğini, itebilirliğini, sıkıştırabilirliğini ve son olarak delebilirliğini tespit eden dört farklı Destek Vektör Makinaları (ing. SVM-Support Vector Machines)[11] eğitimi yapılmıştır. Elde edilen özniteliklerin tek bir vektörde toplanıp tek bir DVM eğitilmesi yerine farklı DVM'lerin eğitilmesinin sebebi 4 farklı sağlarlık sınıfının ortaya çıkaracağı sınıf çokluğudur. Dört farklı sağlarlık sınıfı 7 farklı çeşitten 24 tane el aleti ile eğitilmiştir(Şekil 5).

3.8. Model Testi ve Sonuçları

Eğitilen model, eğitim verisindeki aletlerin çeşitlerinden olmayan 4 alet dahil toplam 11 el aleti ile test edilmiştir. Şekil 6'da toplam 11 el aletinin 44 tane olması gereken sağlarlığı ve DVM'lerin tahmin ettiği sağlarlıklar gösterilmiştir. Şekil 6'da

Şekil 5: Modellerin Eğitildiği El Aletleri

bir aletin sahip olması gereken sağlamlık mavi çember ile, test edilen aletlerin tahmin edilen sağlamlıkları ise siyah tik işareti ile gösterilmiştir. Yukarıdaki sonuçlardan da görülebile-

Aletler	Kesebilen	Delebilen	İtebilen	Sıkıştırabilen
	○		✓	✓
	○	✓		✓
			✓	✓
			✓	✓
		✓		
			✓	
		✓		✓
	✓	✓		✓
			✓	✓
		✓		✓

Şekil 6: Test Edilen Aletlerin Sahip Olması Gereken Sağlamlıkları (Mavi Çember) ve Tahmin Edilen Sağlamlıklar (Siyah Tik)

leceği gibi toplam 44 öznelik içerisinde sadece 3 öznelik yanlış sınıflandırılmıştır. Bu sonuç %93.1 doğruluk payı demektir. Hatalı tahmin edilen 3 tane özneliğin 2 tanesi kesebilmek sınıfına, 1 tanesi ise sıkıştırabilme sınıfına aittir. Kesebilmek sınıfındaki iki hatanın nedeni, kesebilmek özelliğine sahip test edilen iki aletin kesme işleminde ağızlarının tam olarak üst üste gelmemesidir. Eğer eğitim yapılan aletler arasında bu tür ağız tam olarak üst üste gelmeyen ve kesici özelliği bulunan aletler bulunursa, DVM sınıflandırmasını daha uygun yapar ve test edilen kesebilmek özelliğine sahip bu tür aletler de daha yüksek başarı elde edilebilir. Sıkıştırabilme sınıfındaki tek hatanın nedeni ise aynı şekilde eğitim yapılan alet sayısının azlığından ve arada değişebilen aydınlatma sorunundan kaynaklanmaktadır.

4. TEŞEKKÜR

Bu çalışma TÜBİTAK 109E033 no'lu proje tarafından desteklenmiştir.

5. SONUÇ

Sonuç olarak eğitilen modeldeki aletlerin azlığı sonucu çok fazla olmasa da olumsuz bir şekilde etkilemektedir ve aletlerin sadece iki boyutlu incelenmesi sağlamlık sınıflarını kısıtlamaktadır. Buna ek olarak ortamın ışık seviyesindeki değişimler bazı özneliklerin hesaplanmasında değişiklikler yaratmakta ve bu da başarı oranını düşük bir seviyede de olsa etkilemektedir. Daha etkili ve daha geniş bir sağlamlık tespiti ve sağlamlık sınıfı elde etmek için gelecek çalışmalar da etkileşimli görsellikte üçüncü bir boyut kullanılabilir. Bu çalışma insansı robot iCub'ın bu çalışmadaki kadar karmaşık aletler olmasa da daha basit aletler ile etkileşimli görsellik kullanarak alet kullanımını gerçekleştirebileceğini göstermiştir.

6. KAYNAKÇA

- [1] Gibson, J. J. (1986). The ecological approach to visual perception. Original work published 1979. New Jersey: Lawrence Erlbaum Associates.
- [2] Sahin, E., Cakmak, M., Dogar, M. R., Ugur, E., and Ucoluk, G. (2007). To Afford or Not to Afford: A New Formalization of Affordances Toward Affordance-Based Robot Control. *Adaptive Behavior*, 15(4):447-472.
- [3] Shintchi, Y., Sato, Y., Nagai, T. Bayesian Network Model for Object Concept. *Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Proc.* 473-476 2007.
- [4] Nakamura, T., and Nagai, T., "Object Concept Modeling Based on the Relationship among Appearance, Usage and Functions", *Proc. of IROS2010*, pp.4560-4565, Taipei, Oct.2010.
- [5] B. Willimon, S. Birchfield, and I. Walker, "Rigid and non-rigid classification using interactive perception," *Proc. of International Conference on Intelligent Robots and Systems (IROS)*, 2010.
- [6] Dov Katz and Oliver Brock, Manipulating Articulated Objects with Interactive Perception. *Proc. of the IEEE International Conference on Robotics and Automation* 2008.
- [7] Dov Katz and Oliver Brock, Interactive Perception: Closing the Gap Between Action and Perception. *Workshop: From features to actions - Unifying perspectives in computational and robot vision*, *Proc. of IEEE International Conference on Robotics and Automation* 2007.
- [8] B. D. Lucas and T. Kanade (1981), An iterative image registration technique with an application to stereo vision. *Proc. of Imaging Understanding Workshop*, pages 121-130.
- [9] OpenCV: Open Computer Vision Library, <http://opencv.willowgarage.com/wiki/>
- [10] R. Miller and Q.F. Stout, "Efficient parallel convex hull algorithms," *Proc. of IEEE Trans. Comput.*, C-37(12), 1605-1618, 1988.
- [11] V. Vapnik (1998) The support vector method of function estimation. In J. Suykens and J. Vandewalle, ed *Nonlinear Modeling: Advanced Black-Box Techniques*, p55-86, Kluwer Academic Publishers, Boston 1998.