

İnsansı Robot destekli İnteraktif İşaret Dili Eğitimi için İşaret Tanıma

Gesture Recognition for Humanoid Assisted Interactive Sign Language Tutoring

Bekir Sıtkı ERTUĞRUL, Cemal GURPINAR, Hasan KIVRAK, Hatice KOSE
Bilgisayar Mühendisliği Bölümü, Bilgisayar ve Bilişim Fakültesi
İstanbul Teknik Üniversitesi
İstanbul, TÜRKİYE
{bsertugrul, gurpinarcemal, hkivrak, hatice.kose}@itu.edu.tr

Özetçe— Bu çalışma işaret dilini insansı robot destekli taklit bazlı sıralı etkileşim oyunları ile iletişim sorunları yaşayan çocuklara öğretmeyi amaçlayan bir projenin parçasıdır. Bu bildiri etkileşim oyununun otizmlı çocukların da yararlanabileceği şekilde geliştirilmesine yönelik çalışmaları özetler. Otizm Spektrum Bozukluğu (ASD) kısıtlı sosyal etkileşim, iletişim sorunları ve kısıtlı hayal gücü gibi sorunları kapsar. ASD'li çocuklar çoğunlukla robotlara ilgi gösterir. Bu nedenle Robotlar çocuklar ve öğretmenleri/ebeveynleri arasında sosyale etkileşim kurma konusunda destek olarak kullanılabilir. Bu çalışmada bir Nao H25 İnsansı robot öğretmene çocuklara izleme ve taklit yolu ile bazı işaretleri ve basit üst vücut hareketlerini öğretme konusunda yardımcı olmaktadır. İşaret ve hareketleri tanıma konusunda Kinect kamera bazlı bir sistem kullanılmaktadır. Robot bu sistemden gelen tanıma verilerini kullanarak katılımcılara performansları hakkında görsel ve sesli geri bildirimde bulunur.

Anahtar Kelimeler — İnsan-Robot etkileşimi; otizm; taklit oyunları; işaret dili

Abstract— This work is part of an ongoing work for sign language tutoring with imitation based turn-taking and interaction games (iSign) with humanoid robots and children with communication impairments. The paper focuses on the extension of the game, mainly for children with autism. Autism Spektrum Disorder (ASD) involves communication impairments, limited social interaction, and limited imagination. Many such children show interest in robots and find them engaging. Robots can facilitate social interaction between the child and teacher. In this work, a Nao H25 Humanoid robot assisted the human teacher to teach some signs and basic upper torso actions which were observed and imitated by the participants. Kinect camera based system was used to recognize the signs and other actions, and the robot gave visual and audial feedback to the participants based on the performance.

Keywords — Human-robot interaction; autism; imitation games; sign language

I. GİRİŞ

Bu bildiride özetlenen insansı robot ve iletişim problemlili çocuklar ile taklit tabanlı etkileşim oyunu içeren işaret dili eğitimi için devam eden bir çalışmadır. Çalışma özellikle

iletişim sorunu yaşayan otizmlı çocuklar için oyunun uzantısına odaklanmıştır. İletişim insan yaşamı için hayati bir gereksinimdir. Dil edinimi beyin gelişimi ve zeka için ciddi bir süreçtir. İşaret dili işitme engelli ve sözel olarak iletişim kuramayan otistik çocuklar için alternatif bir iletişim yoludur. İşaret dili üst vücut hareketlerini ve yüz hareketlerini içeren görsel bir dildir. Görsel olarak işaret dili tanıyan ve 2 boyutlu araçlarla işaret dili eğitimi veren sistemler mevcuttur [1-4]. İşaret dili gerçeklemede kullanılan robotlar ve robot elleri de literatürde bulunmaktadır [5]. İşaret dili eğitiminde görsel oyunlar da kullanılmaktadır [6-8]

Otizm Spektrum Hastalığı (ASD) sınırlı sosyal etkileşim ve sınırlı hayal gücü ile birlikte iletişim problemlerini içerir. Araştırmacılar ASD'li çocukların tedavi ve rehabilitasyon süreçlerinde robotların kullanılması ile ilgilenmektedirler [9-11]. Otizmlı çocuklar çoğunlukla robotlara ilgi göstermektedirler. Robotlar öğretmen ve çocuk arasındaki etkileşimi kurabilir. Her otizimli çocuk değişik ihtiyaçlara sahiptir. Robot davranışlarının bu süreç esnasında değişime ihtiyacı vardır. Robotlar çocuklar ile etkileşimde başarı ile kullanılmaktadır [12].

Bu oyunda, insansı robot bazı Amerikan ve Türk işaret dilinden oluşan ve çocuğunda robotun hareketlerini taklit edebildiği üst vücut hareketlerini uygular (demo videoları için: <http://humanoid.ce.itu.edu.tr/>) [13,14]. Sonra hareket robot tarafından RGBD kamera kullanılarak tanınır ve robot çocuğu hem sözlü hemde onaylama hareketleriyle harekete geçirir. Hareketler çocuk tarafından robota renkli resimlerin gösterilmesiyle seçilir. Oyun değişik seviyelerin gerektiği üç aşama içermektedir ve oyun değişik hareketler kullanılarak her bir çocuk için bireysel olarak tasarlanır.

II. İŞARET TANIMA

A. El Hareketlerinin Modellenmesi

El hareketi modelleyen bir SMM'in (Saklı Markov Modeli) her saklı durumu bu sembol dizilerinin belirli bir kısmından sorumludur. Homojen SMM'lerde bu bölümlerin süreleri birer geometrik dağılım ile modellenir. Bu süreler her durum için

birbirinden bağımsızdır ve m parametresi modellenmez. Bu kısıtlama, tanınması gereken el hareketi cinsi ve farklı kullanıcı sayısı arttıkça önemli hale gelir. SMM gibi üretici modeller, gözlemler ile durumların ortak olasılıklarını modeller. Ayırıcı modeller ise durumların olasılık dağılımlarını gözlemlere koşullandırır ve gözlemlerin sınıflara bağlı olasılık dağılımını modellemez. Ayırıcı modellerin sınıflandırmada daha etkin olması beklenir. Dolayısıyla el hareketi tanımak için SMM'lerin ayırıcı karşılıkları olan koşullu rasgele alanlar (KRA) ve türevleri de incelenmiştir [15].

KRA'lar bir sınıfın iç dinamiklerini değil, sınıflar arası dinamikleri modeller ve dolayısıyla zaman serisi sınıflandırmaya uygun değildir. Bu nedenle saklı KRA'lar (SKRA) [16] ve gizli dinamik KRA'lar (GDKRA) [17] önerilmiştir. GDKRA sınıfların hem iç hem dış dinamiklerini modellerken SKRA'lar sadece iç dinamikleri modeller ve dolayısıyla izole el hareketi tanıma problemi için daha uygundur.

III. ÖNERİLEN YAKLAŞIM

A. iSign Oyun Uygulaması

Daha önce belirtildiği gibi, otizm ve hesaplamalı yaklaşımlar isimli doktora dersinin çıktısı olan taklit bazlı etkileşim oyunu kinect ve NAO25 insansı robot kullanılarak oluşturulmuştur. Ana amacımız insansı robotun kullanıldığı işaret tabanlı etkileşim oyunlarını otizm terapisi için genişletmektir.

İşaret imitasyon oyunu işaret dili oyununun devamı niteliğindedir. Başlangıç adımı, üst vücut hareketleri kullanılır. Kolları iki yana açık, kollar yukarı ve kollar ileri uzatılmış şekilde basit üst vücut hareketleri (Şekil 1) ile başlayan oyun Amerikan ve Türk işaret dili kelimeleri ile devam eder. Oyunun amacı çocuklara hareketlerin/işaretlerin etkileşim oyunu içinde tanıtılmasını ve taklit edilmesini öğretmektir.



Şekil 1 Oyundan görüntüler.

Bu oyunda hareket gösterici robot olacaktır ve terapist manuel olarak çocuk hareketi yanlış taklit ettiğinde çocuğa yardım edebilecektir.

Oyun 3 aşamadan oluşmaktadır. İlk aşamada çocuk bire bir hareketleri nasıl yapacağını öğrenecektir. Onlar robota hareketin resmini gösterdiğinde, robot hareketi yapıp çocuğun hareketi tekrarlaması için bekleyecektir.

Kinect kamera kullanılarak çocuğun hareketleri değerlendirilebilir ve robota geri bildirimde bulunur. Eğer çocuk hareketi tekrarlarsa robot "hareketi iyi yaptın!" der. Aksi takdirde terapist çocuğa manuel olarak hareketi yapması için yardım eder.

İkinci aşamada oyun bir sportif çalışma gibidir, her hareket resim görüntüsü ve terapist yardımı olmaksızın birkaç kez tekrarlanır. Bu aşamada çocuğun her hareketi öğrendiği kabul edilir.

Üçüncü aşamada oyunu müzikal oyuna dönüştürüyoruz. Robot hareketlere çocuk şarkısı söyleyerek eşlik eder, hareketleri birbiri ardınca dans eder gibi yapar ve çocuğun hareketler dizisini tekrarlaması beklenir.

Robot çocuğun başarı oranını Kinect vasıtası ile anlayarak kayıt edecektir. Ayrıca deneyi yapan terapistin değerlendirilmesi de kayıt edilecektir. Bu oyunlar genellikle terapistle birlikte oynanır. Robot bu oyunlarda yardımcı oyuncu-oyun arkadaşı gibi davranacaktır.

IV. OYUNUN GERÇEKLENMESİ

A. Hareket Tanıma

Hareket tanıma için Kinect kamera ve Microsoft Kinect SDK yazılımı kullanılmıştır. Kinect'in insan iskelet sistemi, eklem pozisyonları algılayabilen özelliği ile insan vücudunun 20 farklı eklem yerlerine ait pozisyon bilgilerine SDK üzerinden erişilmesine olanak sağlar. Daha sonra insanlardan Kinect SDK aracılığıyla alınan üst vücut hareket pozisyon verileri (omuz, dirsek ve bilek koordinatları) çeşitli metodlarla işlenerek yapılan hareket sınıflandırılır.

Başlangıç olarak 3 işaret dili hareketi ve 3 tane temel vücut hareket olmak üzere toplam 6 hareket için 15 farklı insandan (Üniversite öğrencileri) veri toplanmıştır. Her hareket için deneklerden 4 farklı veri alınmıştır. Bu verilerden 2'si eğitim için 2'si de test için kullanılmıştır. Kameradan saniyede 30 frame (30 koordinat bilgisi) kaydedilmiştir. Daha sonra Ayırık Saklı Markov Modeli kullanılarak hareket tanıma çalışması yapılmıştır. Hareketlerin bu şekilde belirlenmesinin sebebi bu hareketlerin otizmliler ve zihinsel özürli çocuklar ile de test edilmesinin amaçlanmasıdır. Bu yüzden çalışmada kullanılmak üzere özel eğitim konusunda uzman terapistler tarafından tavsiye edilen hareketlerin içinden Robot üzerinde de uygulanabilen hareketlerin seçilmesi uygun görülmüştür. Çocukların rahatça uygulayabilmesi, yorulmaması ve sıkılmaması açısından hareket ve tekrar sayıları mümkün olduğunca düşük tutulmuştur.

Öncelikle alınan hareket verisi üzerinde uygulanabilir ayırık gözlem sembolleriyle ve durumlarıyla çalışılabilmesi için bir kümeleme algoritması kullanılmıştır. K-Means metodu kullanılarak hareket verisi ayrıştırılıp kümeleri belirlendikten sonra her hareket için ayrı olmak üzere toplam 6 saklı markov modeli eğitilip model parametreleri Baum-Welch algoritması kullanılarak belirlenmiştir. Daha sonra gelen gözlem dizisinin hangi harekete karşılık geldiğini tespit için sistemin daha önce eğittiği 6 farklı saklı markov modelinin bu gözlemi üretme olasılıkları ileri (forward) algoritması kullanılarak hesaplanmış, en büyük olasılığı sağlayan modele karşılık gelen hareket, bu model ile sınıflandırılmıştır. Eğer eğitim setinde de aynı model ile eğitilmişse hareket doğru tanınmıştır.

Saklı Markov Modelinin karmaşıklığı üzerinde çalışan veriye göre belirlenmelidir. Ses tanıma uygulamalarında kullanılan zaman ilerledikçe bir sonraki duruma geçmeye yada

aynı durumda kalınmasına imkan sağlayan soldan-sağa Saklı Markov Model topolojisi hareket tanıma için de uygun olduğu için tercih edilmiştir. Model karmaşıklığını belirleyen diğer bir parametre durum sayısıdır. Durum dizisi saklı olduğu için en iyi durum sayısı başlangıçta bilinemeyebilir. Eldeki veriye göre deneme yanılma yapılarak seçilebilir. Her bir harekete karşılık gelen hareketin sınıfları kameradan alınan verilere göre belirlendiğinden hepsinin hangi sınıfa ait olduğu bilinmektedir. Bu yöntem bir süpervizörlü öğrenme (denetimli öğrenme) yöntemidir.

B. Projede Kullanılan Robot ve Benzetim yazılımı



Şekil 2. Nao robot bazı TİD ve ASL işaretlerini gerçekleştiriyor

Bu çalışma, işitme engelli kişilere işaret dili öğretiminde robotların kullanımının yaygınlaşmasını hedefleyen projenin bir parçasıdır. Projede el ve parmaklara sahip olduğu için pek çok TİD kelimesini gerçekleştirebilen H-25 Nao robotlar kullanılmaktadır. Ayrıca bu robotların küçük boyutlu olması ve tasarımlarının oyuncaya benzemesi, çocuklar ile çalışırken etkileyici olmaktadır.

Nao robotlar 0.57 m uzunluğunda, 4.5 kg ağırlığındadır. 25 serbestlik derecesine, 500 MHz işlemciye, iki kameraya, dokunma sensörlerine ve sonar sensörlere sahiptir [18]. Projede robotun üst beden eklemleri kullanılmaktadır. Robotun üç parmağının bulunması ve bu parmakların bağımsız hareket edememesi, projede kullanılacak kelimelerin seçimi sırasında en büyük kısıt olmuştur. Çünkü Türk İşaret Dili'nde beş parmağın kullanıldığı kelime sayısı fazladır.

Amaçlanan sistemde, hazırlanan senaryo ile robot işaret dilini öğretecek, bu konudaki başarısı insan eğitmenler ile kıyaslanacaktır. Böylece robot eğitmenlerin de bu alanda çalışabileceği gösterilmiş olacaktır. Robotun kolay güncellenebilir olması ve dış görünüşünün çocukların ilgisini çekmesi, eğitmen olarak kullanılmasında büyük avantajdır.

C. Sonuçlar

6 işarete ait 6 saklı markov model 219 eğitim örneği ile eğitilip, 80 test örneği ile de test edilmiştir. Karşıtlık matrisi(confusion matrix) Çizelge 1'de olarak gösterilmiştir. Bu K ve N değerleri çeşitli değerler ile yapılan testler sonucu en uygun sonucu veren değerlerdir.

Doğru tahmin edilen sınıflar tablonun köşegenlerinde konumlanmıştır. 80 test örneğinin 54 tanesi doğru sınıflandırılmıştır. Hatalı sınıflandırılan gruplar incelediğinde birbirine yakın özellik gruplarından(benzer hareketler) kaynaklandığı gözlemlenebilir. Örnek olarak "masa" işaretini "forward" olarak tahmin etmesi aynı şekilde "baba" ve "up"

işaretlerinin birbirine karıştırması verilebilir. Diğer bir neden ise işaretleri gerçekleştiren kişilerin bazı hareketleri kısmen hatalı veya farklı şekilde gerçekleştirmeleri gösterilebilir.

Çizelge. 1. Sınıflandırma karşıtlık matrisi

		Tahmin Edilen Sınıflandırma					
		Araba (Car)	Baba (Dad)	Forward	Masa (Table)	Side	Up
Gerçek sınıflandırma	Araba (Car)	13	0	0	1	1	0
	Baba(Dad)	0	11	1	0	1	1
	İleri (Forward)	0	0	8	0	0	5
	Masa(Table)	3	1	1	7	0	0
	Yana (Side)	3	1	3	0	6	0
	Yukarı (Up)	1	0	2	1	0	9

İlerleyen çalışmalarda sistem tanıma başarımlarının, saklı koşullu rasgele alanlar, girdi-çıktı saklı markov modelleri ve saklı yarı markov modelleri gibi olasılıksal makine öğrenimi methodları kullanarak artırımı planlanmaktadır.

Robot ve insandan gözlem yapılarak tekrarlanan hareketler arasında istatistik açıdan önemli bir başarımlar farkına rastlanmamıştır.

Oyunun ilk kısmı 15 üniversite öğrencisi ile test edildi. Öğrencilerin 9'u hareketler deneyi yapan araştırmacı tarafından, diğer 6'sına ise Nao H25 tarafından öğretildi. Öğrencilerden 5 basit üst vücut hareketinin yanısıra ASL ve TSL den birkaç kelime (masa, araba, dad, happy) gerçekleştirmesi istendi. Öğrenciler daha öncede işaret dili eğitimi almamışlardı. Zorlukla anlaşılabilen ve taklit edilebilen bazı kelimeler çalışmadan çıkarıldı. Hareketlerin tanınmasında ve taklit edilmesinde henüz ilk çalışmadaki sonuçlar insanla robot eğitmen arasında önemli bir farkın olmadığını göstermektedir. Ayrıca bazı öğrenciler robotun hareketlerini daha basit bulduğunu dolayısıyla tekrar ettikleri her zaman aynı pozisyonda ve hızda olduklarını aksi takdirde öğretmenin hareketlerinin değişebileceğini ifade ettiler. Üniversite öğrencilerinin yanısıra bir normal gelişimli 7 yaş ve 2 otizmlili 6-7 yaş aralığında çocuktan veri alınmıştır. Çocukların veri alımı sırasında sabit durmamaları veri alımını oldukça güçleştirmiştir.

V. SONUÇ

Bu bildiride özetlenen insansı robot ve iletişim problemleri çocuklar ile taklit tabanlı etkileşim oyunu içeren işaret dili eğitimi için devam eden bir çalışmadır. Çalışma özellikle iletişim sorunu yaşayan otizmlili çocuklar için oyunun uzantısına odaklanmıştır.

Bu oyunda, insansı robot bazı Amerikan ve Türk işaret dilinden oluşan ve çocuğunda robotun hareketlerini taklit edebildiği üst vücut hareketlerini uygular. Sonra hareket robot tarafından RGBD kamera kullanılarak tanınır ve robot çocuğu hem sözlü hemde onaylama hareketleriyle harekete geçirir. Hareketler çocuk tarafından robota renkli resimlerin gösterilmesiyle seçilir. Oyun değişik seviyelerin gerektiği üç aşama içermektedir ve oyun değişik hareketler kullanılarak her bir çocuk için bireysel olarak tasarlanır.

REFERANSLAR

- [1] Staner, A. T., and A. Pentland, "Real-Time American Sign Language Recognition from Video using Hidden Markov Models", Technical Report TR-306, Media Lab, MIT.
- [2] Kadous, W., "GRASP: Recognition of Australian Sign Language Using Instrumented Gloves," MSc. Thesis, University of New South Wales, 1995.
- [3] Murakami, K. and H., Taguchi, "Gesture Recognition Using Recurrent Neural Networks," Proceedings of CHI'91 Human Factors in Computing Systems, 1991, pp. 237-242.
- [4] Aran, O., and L. Akarun, A Multi-class Classification Strategy for Fisher Scores: Application to Signer Independent Sign Language Recognition, Pattern Recognition, Vol. 43, no. 5, pp. 1717-1992, May 2010.
- [5] Jaffe, D., Evolution of mechanical fingerspelling hands, for people who are deaf-blind, Journal of Rehabilitation Research and Development, 31, 236-244. 1994
- [6] Adamo-Villani, N., A virtual learning environment for deaf children: Design and evaluation. IJASET – International Journal of Applied Science, Engineering, and Technology, 16, 18-23, 2006
- [7] Lee, S., V. Henderson, H. Hamilton, T. Starner, H. Brashear, and S. Hamilton, "A Gesture-based American Sign Language (ASL) Tutor for Deaf Children", Proceedings of CHI (Computer-Human Interaction). Portland, OR. April 2005
- [8] Greenbacker, C., and K. McCoy, The ICICLE Project: An Overview. First Annual Computer Science Research Day, Department of Computer & Information Sciences, University of Delaware, Feb 2008.
- [9] Billard A., Robins B., Nadel J., Dautenhahn K., Building Robota, a Mini-Humaooid Robot for the Rehabilitation of Children with Autism. RESNA Assistive Technology Journal, vol.19, 2006
- [10] Dautenhahn K., Roles and Functions of Robots in Human Society - Implications from Research in Autism Therapy. Robotica 21(4), pp. 443-452, 2003
- [11] Kozima H., C. Nakagawa, and Y. Yasuda. Children-robot interaction: a pilot study in autism therapy. Prog Brain Res, 164:385-400, 2007.
- [12] Kose-Bagci, H, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Drum-mate: interaction dynamics and gestures in human-humanoid drumming experiments," Connection Science, vol. 22, no. 2, pp. 103- 134, 2010.
- [13] Kose, H., R. Yorganci , H. E. Algan, and D.S. Syrdal, "Evaluation of the Robot Assisted Sign Language Tutoring using video-based studies", SORO special issue on "Measuring Human-Robot Interaction, 2012, DOI: 10.1007/s12369-012-0142-2
- [14] Kose, H., and R. Yorganci , "Tale of a robot: Humanoid Robot Assisted Sign Language Tutoring", 11th IEEE-RAS International Conference on Humanoid Robots, Bled, Slovenia (HUMANOIDS 2011) , pp 105 – 111, 2011,
- [15] Lafferty J. D., A. McCallum, ve F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in Proceedings of the Eighteenth International Conference on Machine Learning, pp. 282-289, 2001
- [16] Wang S. B., A. Quattoni, L. Morency, ve D. Demirdjian, "Hidden conditional random fields for gesture recognition," Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, USA, pp. 1521-1527, 2006
- [17] Morency L., A. Quattoni, ve T. Darrell, "Latent-dynamic discriminative models for continuous gesture recognition," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 0, pp. 1-8, 2007.
- [18] Aldebaran Robotics Choregraphe. [Online]. <http://www.aldebaran-robotics.com/en/> , Available: 05.04.2013