

ENGELLİLER İÇİN BİLGİSAYAR UYGULAMALARI VE İŞARET DİLİ EĞİTİMİ

Oya ARAN, Bogazici Üni. Bilgisayar Müh. Böl., Bebek 34342-Istanbul, aranoya@boun.edu.tr

Cem KESKİN, Bogazici Üni. Bilgisayar Müh. Böl., Bebek 34342-Istanbul, keskin@cmpe.boun.edu.tr

Lale AKARUN, Bogazici Üni. Bilgisayar Müh. Böl., Bebek 34342-Istanbul, akarun@boun.edu.tr

Özet

Bu çalışmada, gerçek zamanlı el izleme ve Saklı Markov Modelleri (HMM) kullanarak el hareketlerini tanımaya dayalı bir etkileşimli arayüz geliştirilmiş ve bu sistem İşaret Dili Eğitimi için uyarlanmıştır. Etkileşimli hareket eğitimi ve kamera kalibrasyonu için birer yazılım aracı da içeren bu sistem, değişik kişisel bilgisayar uygulamalarında kullanılabilir, özel donanım gerektirmeyen, basit ve gülbüz bir arayüz örneği olarak önerilmektedir. Önerdiğimiz sistem renkli eldiven giymiş bir kullanıcının el koordinatlarını stereo kameralar aracılığıyla bulup yapılan üç boyutlu el hareketlerini tanımaktadır. Tanıma aşamasında Saklı Markov Modelleri kullanılmış, bu modeller Türk İşaret Dili'ndeki işaretlerden oluşan eğitim kümesi ile eğitilmiş ve eğitilen modeller İşaret Dili Eğitimi programı tarafından kullanılmıştır. Sistem, yedi farklı işaret tipinden oluşan veri kümesi üzerinde %95 başarı göstermiştir.

Anahtar Kelimeler: El izleme, işaret dili tanıma, Saklı Markov Modelleri, Kalman Süzgeci

COMPUTER APPLICATIONS FOR DISABLED PEOPLE AND SIGN LANGUAGE TUTORING

Abstract

In this work, we developed an interactive interface with real time hand tracking and hand gesture recognition based on HMMs. The system is also adapted to a Sign Language Tutoring application. The system requires no special hardware and it is simple and robust with included tools for interactive training and interactive camera calibration. Proposed system finds the 3D hand coordinates of a user with a colored glove using 2 web cameras and recognizes the performed gesture or sign. Hidden Markov Models are trained with a database of signs from Turkish Sign Language and these models are used by the Sign Language Tutoring application. The performance of the system is 95% for the database of seven sign classes.

Keywords: Hand tracking, sign language recognition, Hidden Markov Models, Kalman Filter

1. Giriş

İnsan-bilgisayar etkileşiminde amaçlanan doğallık ve rahatlığı sağlamak için, el hareketlerini görsel olarak tanıyan sistemler dikkate değer bir alternatiftir. Bu sistemler aynı zamanda işaret dili tanıma sistemlerinin de temelini oluşturur. Günümüzün gelişmiş kişisel bilgisayarları basit kameralar yardımıyla üç boyutlu el hareketlerini yorumlayacak ve işaret dillerini tanıyabilecek kapasiteye ulaşmıştır.

1.1. İnsan Bilgisayar Etkileşimi ve El hareketleri Tabanlı Sistemler

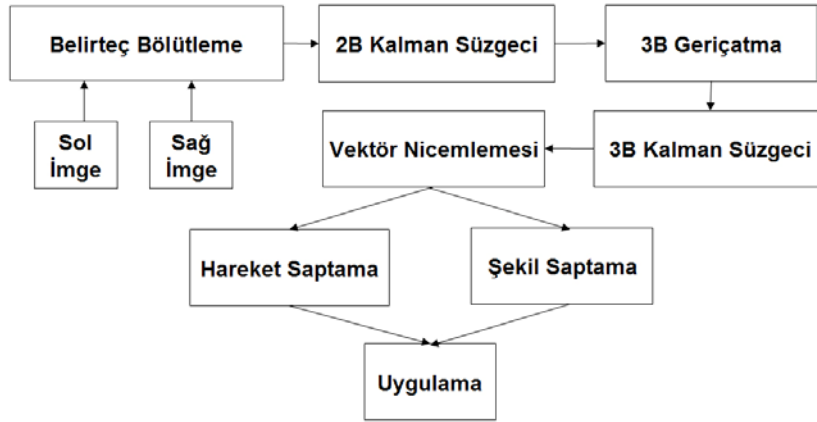
El hareketlerinin uzay-zamansal değişkenlik göstermesi ve ardışık hareketlerin ayrılmasının zorluğu el hareketlerinin tanınmasını karmaşık bir problem haline getirir. Bu konuda Saklı Markov Modelleri (Hidden Markov Models - HMM) değişik uygulamalarda başarılı olmuş ve yaygın kullanım alanı bulmuştur [1].

HMM kullanılan ilk dinamik hareket tanıma sistemlerinden biri Amerikan işaret dilini tanımak üzere Starner ve Pentland tarafından geliştirildi [3]. Oka, Sato ve Koike de yine HMM kullanarak iki boyutlu el hareketlerini tanıyan görmeye dayalı sistemlerinde gürültünün etkisini azaltmak için Kalman süzgecinden yararlandılar [4]. Bu araştırmacılar daha önceki bir çalışmalarında üç boyutlu el hareketlerini yapay sinir ağları kullanarak tanımaya yönelik bir sistem de geliştirmişlerdir [5].

Bu çalışmada, gerçek zamanlı el izleme ve üç boyutlu el hareketlerini tanımaya dayalı etkileşimli bir arayüz ve bu arayüzü kullanan bir işaret dili eğitmeni tasarladık. Önerdiğimiz sistemde üçüncü boyutun

algılanması için iki adet renkli kameradan yararlanıyoruz. Ayrıca gürültünün ayıklanması için 2B ve 3B Kalman süzgeci, el hareketlerinin tanınması içinse HMM kullanıyoruz.

Sistemin işleyişi Şekil 1'de görülmektedir.



Şekil 1 Sistemin Akış Çizeneği

1.2. İşaret Dili ve İşaret Dili Tanıma

İşaret dili, temelinde el hareketleri ve el şekline dayanan fakat bunların yanında yüz mimiklerinin, baş ve vücut hareketlerinin de kullanıldığı görsel bir dildir. İşaret dilleri, işaretlerin basit bir bileşimi olmaktan çok, aynı sözlü diller gibi karmaşık gramer yapısına sahiptir ve her ülke için farklıdır. İşaret dillerinin çevresinde konuşulan sözlü dille benzerlik göstermesi gerekmez ama etkiler görülebilir. Kökleri Osmanlı işaret diline kadar uzandığı düşünülen Türk İşaret Dili hakkında çok fazla kaynak yoktur. İşitme engelliler okullarında işaret dili eğitiminin 1953 yılında yasaklanmasıyla beraber Türk İşaret Dili kesintiye uğramış ve Türkiye'nin farklı yerlerinde konuşulan işaret dilleri birbirinden çeşitli farklılıklar göstermeye başlamıştır. Türk İşaret Dili hakkında hazırlanmış en kapsamlı kaynaklar, Milli Eğitim Bakanlığı tarafından 1995 yılında çıkarılmış olan görsel bir kitap [6] ve Türkiye Bilimler Akademisi ve Koç Üniversitesi tarafından hazırlanan İnternet sitesidir (<http://turkisaretdili.ku.edu.tr/>).

İşaret dilli tanıma problemi, görsel dili oluşturan tüm öğelerin bilgisayarda analiz edilmesi ve yapılan tek bir işaretin ya da genel olarak işaret dili ile yapılan bir konuşmanın anlaşılması olarak tanımlanabilir. Bu tür bir sistemin geliştirilmesi, işitme ve konuşma engelliler ve engelli olmayanlar arasındaki iletişimi kolaylaştırması açısından önemlidir. Kamera görüntülerine dayanan kapsamlı bir işaret dili tanıma sistemi aşağıdaki problemleri çözmelidir.

- Kamera görüntüsünde el, yüz, baş ve vücudun çeşitli bölümlerinin bulunması ve izlenmesi
- El hareketinin tanınması
- El şeklinin tanınması
- Yüz ifadesinin tanınması
- Baş hareketlerinin tanınması
- Farklı öğelerden gelen bilgilerin birleştirilerek yapılan işaretin tanınması
- Arka arkaya gelen işaretlerin tanınması
- Dilin gramer kuralları kullanılarak cümlenin tanınması

Bu problemler, örüntü tanıma, bilgisayarda görme, doğal dil işleme, psikoloji ve dil bilim alanlarında ortak araştırma ve çalışma yapılarak çözülebilir. Dolayısıyla, işaret dili tanıma, bir çok araştırma alanda ortak çalışma yapmayı gerektiren karmaşık bir problemdir.

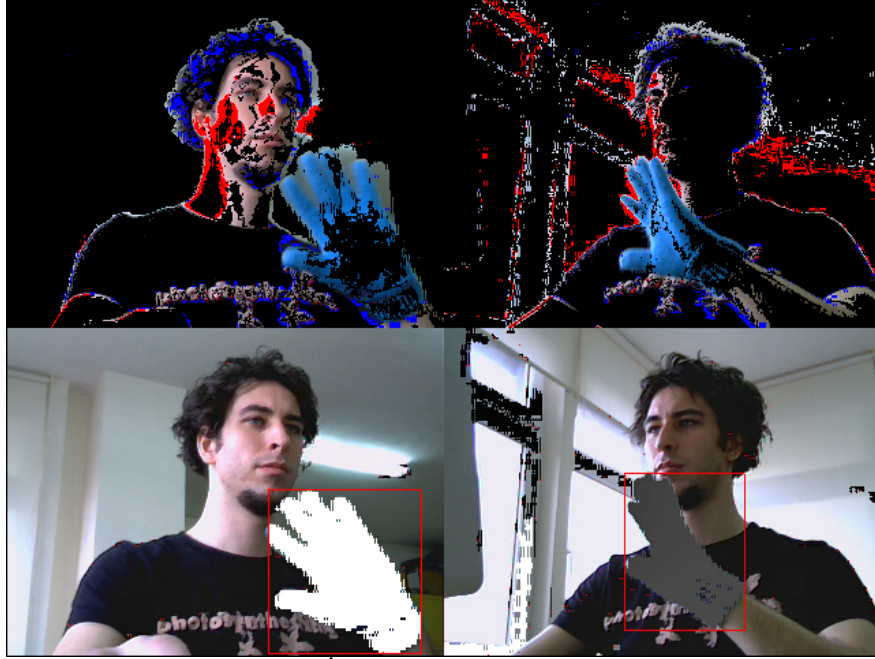
2. El İzleme

Sistemimizin altyapısı el izleme ve hareket tanıma olarak iki ana kısımda incelenebilir. El izleme modülü sırasıyla kamera kalibrasyonu, belirtecin sisteme tanıtılması, kullanıcının elinin imgelerde bulunması, geriçatma ile elin üç boyutlu gezinesinin çıkarılması ve gezinedeki gürültünün etkisinin yok edilmesi aşamalarından oluşur.

2.1 Belirtecin Tanıtılması

Vücudun görüntüye giren diğer kısımlarından dolayı elin ayırılması karmaşık bir problemdir. Belirteç kullanımı bu problemi büyük oranda kolaylaştırdığı için karmaşık arkaplanlar ve değişken ışık koşullarında elin tanınabilmesi amacıyla belirteçlerden yararlanılmaktadır. Belirteç kullanımı el izleme aşamasında elin görüntülerden renk bölütlemesi yoluyla çıkarılmasını sağlar. Dinamik ışıklandırma koşullarında renk tonu (hue) ve renk doygunluğu RGB değerlerine göre daha az değişim gösterdiğinden renk bölütlemesi için bu değerleri kullanılmaktadır. Ancak renk tonu gri renkleri için tanımlı olmadığı için ve griye yakın renkleri için de kullanışlı olmadığı için çok aydınlık, çok karanlık ve griye çok yakın değerleri renk bölütlemesinde kullanılmıyoruz.

Seçilen belirtecin ten rengiyle karışmayan düz renkli bir eldiven olması gerekir. Renk önceden belirlenmiş olmadığı için ve ışık şartlarına göre değişiklik gösterdiği için seçilen belirtecin sisteme tanıtılması gerekmektedir. Tanıtılma işlemi için sistemin belirteç ayırma programı çalıştırılıp yalnızca eldiven hareket ettirilir ve her iki kameradan ardışık ikişer imge yakalanır. Ardından bu iki imgenin karşılıklı piksellerinin RGB değerleri karşılaştırılır ve farkları bir eşik değerini geçmeyen pikseller ilk imgeden silinir. Buna ek olarak griye yakın değerler, çok aydınlık ve çok karanlık bölgeler işaretlenir ve tanıma işleminde kullanılmaz (Şekil 2).



Şekil 2 Fark İmgelerinden Belirtecin Tanınması

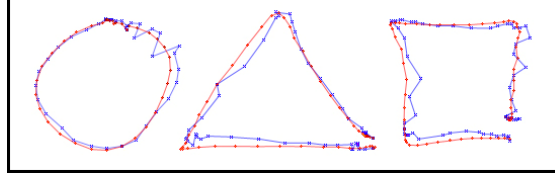
Hareket esnasında eldiven dışındaki nesnelere sabit tutulduğundan oluşturulan imgede kalan en büyük alana sahip bölge belirteceye ait olacaktır. Bu bölgenin piksellerinin HSV renk sistemindeki renk tonu ve doygunluğu histogramları çıkarılır ve en fazla piksel içeren aralıktaki piksellerin ortalamaları bulunur. Bu değerler sistemin geri kalanında el izleme aşamasında kullanılır.

2.2 Belirtecin İmgelerde Bulunması

Belirteç, görüntü dizisindeki imgelerden bağlantılı bileşen algoritması kullanılarak çıkarılır. Bağlantılı bileşen algoritması birleştirme şartı olarak pikselin renk tonuna, renk doygunluğuna, pikselin çok aydınlık veya karanlık olmamasına ve griye çok yakın olmamasına bakar. Bu işlem sırasında bağlanan bileşenlerin renk tonunun belirtecininkine olabildiğince yakın olması için çift eşik yöntemi kullanılır. Kullanılan tüm parametreler uygulama üzerinden kullanıcı tarafından değiştirilebilir. Bulunan bileşenlerin piksel alanları bileşen bağlama işlemi esnasında hesaplanır ve görüntü çözünürlüğüne bağlı dinamik bir eşik değerinden küçük olmamak kaydıyla en büyük alana sahip bileşen el olarak ayırılır. Elin sınırlayıcı kutusunun uzanım özellikleri sonraki aşamalarda elin genel şeklinin tanınmasında kullanılır.

2.3 Gezinenin Yumuşatılması

Elin ardışık imgelerden bulunan koordinatları kullanılarak elin izlediği gezege bulunur. Ancak bu gezege örnekleme gürültüsü yüzünden tanıma için elverişli değildir. Örnekleme çok kısa zaman aralıklarıyla yapıldığı için, belirtecin ardışık görüntüler arasında sabit hızla doğrusal hareket ettiğini ve gezinmesinin rastlantısal sarsımlara uğradığını varsayıyoruz. Bu varsayım üzerine kurulmuş bir Kalman Süzgeci kullanarak ayrı kameralardan gelen iki boyutlu gezinmeleri yumuşatıyoruz. Şekil 3'te Kalman Süzgeci'nin örnekleme ile bulunmuş asıl koordinatlar üzerindeki etkisi verilmiştir.



Şekil 3 Kalman Süzgecinin Etkisi

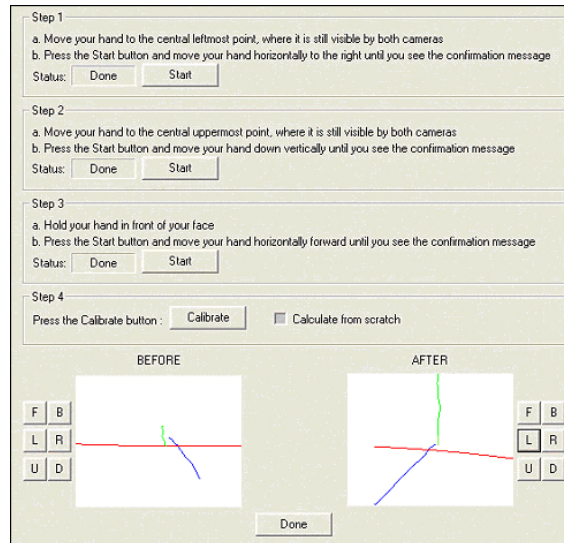
2.4 Geriçatma İşlemi

Parmak ucunun üç boyutlu koordinatlarının bulunduğu geriçatma işlemi iki kameranın kalibrasyon matrisleri ile en küçük kareler yöntemi kullanılarak yapılmaktadır. Bu sebeple iki kameranın da kalibrasyon matrislerinin önceden biliniyor olması gereklidir. Uygulama kullanıcıya iki farklı kamera kalibrasyonu yöntemi sağlamaktadır. İlk yöntem kalibrasyon işlemi kolaylaştırmak amacıyla Şekil 4'te görülen kalibrasyon nesnesinin üzerindeki noktaları otomatik olarak bulup eşleştirerek obje koordinatları ile imge koordinatları arasındaki ilişkiyi kalibrasyon matrislerini çıkarır.



Şekil 4 Kalibrasyon Objesi

İkinci kalibrasyon yöntemi özel bir nesne kullanmak yerine kullanıcının eliyle çizdiği birbirine dik üç çizginin iki imge üzerindeki izdüşümlerinden temel görüntü matrislerini ve kalibrasyon matrisini çıkarır. Aynı araç var olan kalibrasyon matrislerini test etmekte ve açılarını düzeltmekte de kullanılabilir.



Şekil 5 El İzleme Yöntemi İle Kamera Kalibrasyonu

Kalibrasyon matrisleri kameralara ve konumlarına bağlıdır ve bir kez bulunduktan sonra büyük bir değişiklik olmadığı sürece kalibrasyon işlemini tekrarlamak gerekmez. Ancak ölçüm gürültüsü rasgele olduğu için yalnızca imge koordinatlarının filtrelenmesi geriçatma işleminde tatmin edici sonuç vermez. Oluşturulan 3B gezingedeki sarsımları azaltmak için gezinge tekrar Kalman süzgecinden geçirilir. İkinci filtreleme işlemi sistemin tanıma başarısını önemli ölçüde arttırmaktadır.

3. Hareket Tanıma

Hareket tanıma modülü yeni hareketlerin tanınması için gerekli eğitimi, gezinenin tanınması için gerekli niceleme işlemi, hareketi tanıma ve el şeklini tanıma aşamalarını kapsar. Sistemde tanımlı her el hareketi için birer HMM ve el şekli tanıma kuralı bulunmaktadır. Bir dizinin tanınması için dizi varolan tüm HMM'lere girdi olarak verilir. Dizinin anlamlı el hareketlerinden biri olarak kabul edilebilmesi, bu modellerden birinin verdiği olabirliğin diğerlerinininkileri geçmesi ile mümkündür. Ancak kullanıcının sadece anlamlı olan hareketleri yapma kısıtlaması olmadığından dolayı anlamlı hareketleri anlamsızlardan ayırdedebilmek için bu koşul yeterli değildir. Aday modelin olabirliğinin bir eşik değerini de geçmesi gerekmektedir. HMM'in değişik uzunluk ve karmaşıklıkta el hareketlerine verdiği tanıma olabirlikleri büyük farklılıklar gösterdiği için sabit bir eşik değeri belirlemek iyi sonuç vermez. Dolayısıyla Lee ve Kim tarafından önerilmiş olan uyarlanı eşik modeline benzer bir model oluşturduk [2]. Bir hareketin tanınmasının ilk şartı, olabirliğinin diğer hareketlerle beraber eşik modelinin olabirliğini de geçmesidir (Şekil 6). Daha sonra sistem tanımlı şekil kuralları ile elin basit şekilsel özelliklerine bakarak hareketin ne olduğuna karar verir.

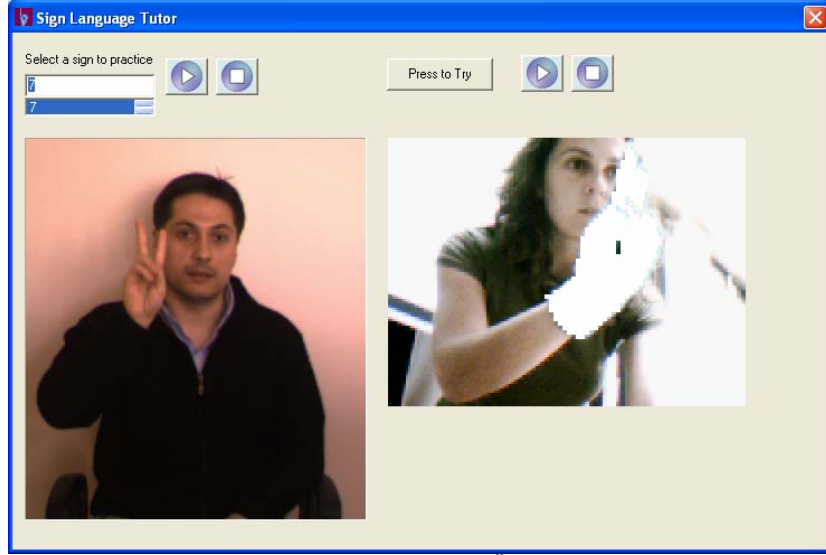
Sisteme tanımlanan hareketlerin süresi önceden belirli değildir. Dolayısıyla gelen her yeni girdi olası bir bitiş noktası olarak varsayılır ve sınırları uygulama üzerinden değiştirilebilir bir aralık taranarak olası başlangıç noktaları aranır. Bu aralıktan başlayan tüm diziler için olabirlikler hesaplanır ve eğer bir dizi için HMM modellerinden birinin verdiği olabirlik eşik modelin verdiği olabirliği aşarsa söz konusu hareket tanınmış olarak kabul edilir.

4. Bilgisayarda İşaret Dili Eğitimi

İşaret dili öğreniminde öğrenilen işaretlerin doğru yapıp yapılmadığının geçerlenmesi ve işaretin tekrarlanması önemlidir. Bu çalışmada geliştirilen program, bu geçerlemenin bir bilgisayar programı ile yapılmasını hedefler. İşaret dili, bir çok görsel öğeyi kullanan karmaşık bir problem olmasına karşın bu çalışmada sadece işaret dilinin en önemli öğesi olan el hareketleri analiz edilmiştir. Programın gereksinimleri, kısıtlamaları şu şekildedir:

- Programın kullandığı veri tabanı Türk İşaret Dili'nden 7 işareti içerir. Programın arayüzü kullanılarak, işaret dili bilen bir kullanıcı tarafından yeni işaretler eklenebilir, bu şekilde veri tabanı genişletilebilir.
- El şekli ile ilgili çok basit özellikler kullanılmakla beraber, programda el hareketlerinin tanınması üzerine yoğunlaşmıştır. Yüz ifadeleri, baş ve vücut hareketleri kullanılmamıştır.
- Program tek başına yapılan işaretleri tanımayı amaçlar. Arka arkaya yapılan işaretler bu çalışmanın dışındadır.
- Program gerçek zamanlı çalışır.
- Kullanıcının tek renkli bir eldiven giymesi gerekmektedir. Eldivenin rengi çevredeki nesnelere farklı olmalıdır.
- Çevredeki ışığın yetersiz olması elin izlenmesinde sorunlar yaratabilir.

"İşaret Dili Eğitmeni" programı yukarıda anlatılan el izleme ve hareket tanıma tekniklerini kullanır. Programın arayüzü iki aşamadan oluşur: Öğrenme aşaması ve uygulama aşaması. Öğrenme aşamasında kullanıcı önce bir işaret seçer. Seçilen işaret için daha önceden kaydedilmiş bir video kullanıcıya gösterilir. Kullanıcı bu videoyu, uygulama aşamasına geçmeye hazır olana kadar tekrar tekrar izleyebilir. Uygulama aşamasında, kullanıcıdan, seçmiş olduğu işareti yapması istenir. Kullanıcının yaptığı işaret analiz edilir. Yapılan analize göre kullanıcıya yaptığı işaretin doğru olup olmadığı bildirilir. Kullanıcı ayrıca yaptığı hareketin videosunu izleyebilir.



Şekil 7 İşaret Dili Eğitmeni Arayüzü – Öğrenme ve uygulama aşamaları

4. Deneyler

Deneylerde Türk İşaret Dili'nden seçilen 7 işaret için, 7 farklı kullanıcıdan alınan veriler kullanılmıştır. Her kullanıcı 7 değişik işareti en az 15 kere tekrarlamıştır. 4 kullanıcının verisi eğitim kümesinde kullanılmış, kalan 3 kullanıcının verisi ise sınav kümesinde kullanılmıştır. Kullanılan işaretlerin özellikleri Tablo 1'de gösterilmiştir.

İşaret	Hareketin özelliği	İşaretin yapıldığı yer	Temel Uzanım Yönü
7	Dikey hareket	Başın yanı	Y ekseninde
anne	Yatay hareket	Göğüs	X ekseninde
cuma	Dikey hareket	Yüz	X ekseninde
görmek	Z ekseninde hareket	Gözler	Y ekseninde
kolay	Yatay hareket	Ağız	X ve Y ekseninde eşit
konuşmak	Z ekseninde dairesel hareket	Ağız	X ekseninde
yok	Z ekseninde yarım dairesel hareket	Omuz üstü	Y ekseninde

Tablo 1 Kullanılan işaretler ve özellikleri

Her işaret sınıfı için Saklı Markov Modelleri eğitilmiş ve bu modellerin durum sayılarının belirlenmesi otomatik olarak yapılmıştır. Önceden belirlenmiş bir sayıyla başlanarak modeller eğitilmiş ve bir sonraki duruma geçme olasılığı belli bir değerin üstünde olan durumlar modelden çıkarılmıştır. Bu yöntem ile yaklaşık 4-5 durumlu modeller elde edilmiştir. Sistemin başarısı Tablo 2'de gösterilmiştir. Bazı işaretlerin diğerlerine göre daha yüksek başarıyla tanındığı görülmektedir. Bazı işaretler ise birbirleriyle karıştırılmaktadır. El şekli ve elin yeri ile ilgili daha fazla bilgi kullanıldığı takdirde sistemin başarısı daha da artacaktır.

İşaret	Örnek Sayısı			Başarı (%)	
	Toplam	Eğitim Kümesi	Sınav Kümesi	Eğitim Kümesi	Sınav Kümesi
yedi	306	184	122	83.7	87.7
anne	211	127	84	77.9	86.9
cuma	137	82	55	100	100
görmek	101	61	40	100	100
kolay	124	74	50	98.6	100
konuşmak	96	58	38	91.4	100
yok	102	61	41	82	90.2
toplam	1077	647	430	90.5	95.0

Tablo 2 Tanıma sonuçları

5. Sonuç

Bu çalışmada üç boyutlu el hareketlerinin gerçek zamanlı takibi ve işaret dilinin tanınması üzerinde çalıştık. Öncelikle belirteç çeşitli görüntü işleme algoritmaları kullanılarak kameralardan gelen görüntülerde ayırıldı. Ardından her iki kameradan gelen görüntüler kullanılarak üç boyutlu geriçatma işlemi uygulandı. Daha sonra elde edilen gezinge önceden eğitilmiş HMM'lere gönderildi ve olabilirliği eşik modelini geçen hareketler tanındı. İşaret dili eğitim uygulamasında hareketlerin yanısıra genel el şekil bilgisi de dikkate alındı.

Geliştirilen bu program işitme ve konuşma engelliler tarafından kullanılabilmesi gibi engelli olmayanlar tarafından da kullanılabilir ve kullanıcılara işaret dili öğretmeyi hedefler. Uygulamanın geliştirilmesi için, Türk İşaret Dili'ndeki tüm işaretlerin bulunduğu bir veri kümesi hazırlanarak programa entegre edilmesi gerekmektedir.

Bu çalışma DPT/03K120250 "Algısal İnsan Bilgisayar Etkileşimi" ve AB 6. çerçeve programı SIMILAR projeleri tarafından desteklenmektedir.

Kaynaklar

- [1] RABINER, L.R., 1989, A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proc IEEE, vol.77, pp.257-285.
- [2] LEE H., KIM J.H., 1999, An HMM-Based Threshold Model Approach for Gesture Recognition, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 961-973.
- [3] STRANER T., PENTLAND A., 1995, Real Time American Sign Language Recognition from Video Using Hidden Markov Models, Technical Report TR-375, MIT's Media Lab.
- [4] OKA, K., SATO, Y., KOIKE, H., 2002, Real-Time Fingertip Tracking and Gesture Recognition, IEEE Computer Graphics and Applications, November-December, 2002, pp. 64-71.
- [5] SATO Y., SAITO, M., KOIKE, H., 2001, "Real-time input of 3D pose and gestures of a user's hand and its applications for HCI," Proc. 2001 IEEE Virtual Reality Conference (IEEE VR2001), pp. 79-86.
- [6] MEB Yayınları, 1995, Yetişkinler için İşaret Dili Kılavuzu, T.C. Milli Eğitim Bakanlığı Özel Eğitim Rehberlik ve Danışma Hizmetleri Genel Müdürlüğü, Ankara
- [7] KESKİN, C., ERKAN, A.N., AKARUN, L., 2002, Real time hand tracking and 3D gesture recognition for interactive interfaces using HMM, Proceedings of ICML, Istanbul, Turkey