

ETKİLEŞİMLİ ARAYÜZLER İÇİN GERÇEK ZAMANLI EL İZLEME VE HMM TABANLI ÜÇ BOYUTLU HAREKET TANIMA

A. N. Erkan, C. Keskin, L. Akarun
Bilgisayar Mühendisliği Bölümü
Boğaziçi Üniversitesi
Tel: (212) 358 15 40 / 1858
Fax: (212) 287 24 61

E-Mail: ckeskin@ttnet.net.tr, {erkanays,akarun}@boun.edu.tr

Bildiri Konusu : Çoklu-Ortam ve İnsan-Makine Etkileşimi

Özetçe

Bu çalışmada gerçek zamanlı el izleme ve Saklı Markov Modelleri (HMM) kullanarak el hareketlerini tanımaya dayalı bir etkileşimli arayüz tasarladık. Önerdiğimiz sistem renkli eldiven giymiş bir kullanıcının el koordinatlarını stereo kameralar aracılığıyla bulup yapılan üç boyutlu el hareketlerini tanımaktadır.

Etkileşimli hareket eğitimi ve kamera kalibrasyonu için birer yazılım aracı da içeren bu model, değişik kişisel bilgisayar uygulamalarında kullanılabilir, özel donanım gerektirmeyen, basit ve gürbüz bir arayüz örneği olarak önerilmektedir. Sistem bir çift sıradan webcam ile toplam sekiz hareketi tanımada 160 denemede %98.75 başarı göstermiştir.

1. Giriş

İnsan-bilgisayar etkileşiminde amaçlanan doğallık ve rahatlığı sağlamak için, el hareketlerini görsel olarak tanıyan sistemler dikkate değer bir alternatiftir. El hareketlerinin uzay-zamansal değişiklik göstermesi ve ardışık hareketlerin ayrılmasının zorluğu el hareketlerinin tanınmasını karmaşık bir problem haline getirir. Bu konuda 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 tasarladık. Önerdiğimiz sistemde üçüncü boyutun algılanması için iki adet renkli kameralardan yararlanıyoruz. Ayrıca gürültünün ayıklanması için

2B ve 3B Kalman süzgeci, el hareketlerinin tanınması içinse HMM kullanıyoruz.

2. El İzleme ve Hareket Tanıma

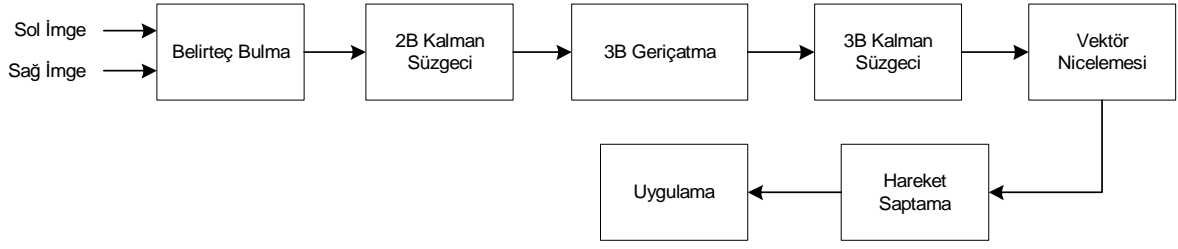
Sistemimiz bilgisayarla üç boyutlu görü ve hareket tanıma olarak iki ana kısımda incelenebilir. Üç boyutlu görme modülünün önemli aşamaları öncelikle elin görüntülerde ayırılması, ardından izlenen parmağın üç boyutlu koordinatlarının edinimi ve sistemdeki gürültünün etkilerinin ayıklanmasıdır. Hareket tanıma modülü ise yeni hareketlerin tanınması için gerekli eğitimi ve hareketi tanıma aşamalarını kapsar. Tanıma için gerekli girdi görü birimi tarafından oluşturulduğu için görü birimi program akışında hareket tanımadan önce gelir. Sistemin işleyişi Şekil 1'de görülmektedir.

2.1. Üç Boyutlu El İzleme

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 yararlanmaktayız.

Belirteç kullanımı el izleme aşamasında elin görüntülerden renk bölütlemesi yoluyla çıkarılmasını sağlar. Dinamik ışıklandırma koşullarında HSV renk koordinat sistemindeki renk tonu H, RGB değerlerine göre daha az değişim gösterdiğinden renk bölütlemesi için renk tonu kullanılmaktadır.

Kullanıcı belirteç olarak ten rengiyle karışmayan düz renkli bir eldiven seçebilir. Belirteç ayırtma programı çalıştırılıp yalnızca eldiven hareket ettirilir ve her iki kameralardan ardışık ikişer imge yakalanır. Ardından bu iki imgenin karşılıklı piksellerinin RGB değerleri karşılaştırılır ve farkları bir eşik değerini geçmeyen pikseller ilk imgeden silinir. Hareket esnasında eldiven dışındaki nesnelere sabit tutulduğundan oluşturulan imgede kalan en büyük alana sahip bölge belirtece ait olacaktır. Bu



Şekil 1 Sistemin Akış Çizeneği

bölgenin piksellerinin HSV renk sistemindeki renk tonu ortalamaları bulunur ve daha sonra el izleme aşamasında kullanılır.

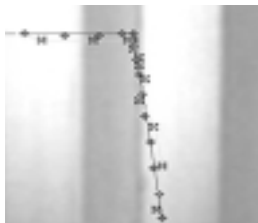
Belirteç, görüntü dizisindeki imgelerden bağlantılı bileşen algoritması kullanılarak çıkarılır. Bu işlem sırasında bağlanan bileşenlerin renk tonunun belirtecinkine olabildiğince yakın olması için çift eşik yöntemi kullanılır. Bulunan bileşenlerin piksel alanları bağlantılı bileşenler algoritması ile hesaplanır ve toplam görüntü çözünürlüğüne bağlı dinamik bir eşik değerinden küçük bileşenler ayıklanır. Daha sonra geri kalan bileşenlerin sınırlayıcı kutularının en/boy değerlerine bağlı olarak el ayır edilir.



Şekil 2 El Kipleri

Programın akışını kontrol edebilmek amacıyla Şekil 2'deki el kiplerinden yararlanıyoruz. İzlenen parmak kullanılan el kipine göre değiştiği için sınırlayıcı kutunun boyutlarına bakarak elin kipini, bu bilgiyi kullanarak da parmağın yerini buluyoruz.

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 Kalman Süzgeci

Parmak ucunun üç boyutlu koordinatları en küçük kareler bazlı bir yöntem kullanılarak bulunmaktadır. Bu yöntem için iki kameranın da kalibrasyon matrisi gereklidir. Kalibrasyon işlemini

kolaylaştırmak amacıyla sisteme, Şekil 4' te görülen kalibrasyon nesnesi için özelleşmiş, bu nesne üzerindeki noktaları otomatik olarak bulup eşleştiren bir yazılım aracı ekledik.



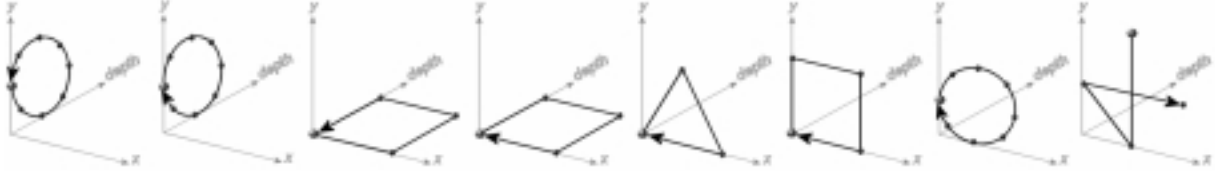
Şekil 4 Kalibrasyon Objesi

Ölçüm gürültüsü rasgele olduğu için yalnızca imge koordinatlarının filtrelenmesi geriçatma işleminde tatmin edici sonuç vermez. Oluşturulan 3B gezinmedeki sarsımları azaltmak için gezinme tekrar Kalman süzgecinden geçirilir. İkinci filtreleme işlemi sistemin tanıma başarısını önemli ölçüde arttırmaktadır.

2.2 Hareket Tanıma

Sistemde tanımlı her el hareketi için birer HMM 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 olabilirliğin diğerlerinininkileri geçmesi ile mümkündür. Bunun yanı sıra, aday modelin olabilirliğinin bir eşik değerini de geçmesi gerekmektedir. HMM'in değişik uzunluk ve karmaşıklıkta el hareketlerine verdiği tanıma olabilirlikleri büyük farklılıklar gösterdiği için sabit bir eşik değeri belirlemek iyi sonuç vermez. Dolayısıyla [2]'de önerilmiş olan uyarlanı eşik modeline benzer bir model oluşturduk.

Herhangi bir el hareketinin uzunluğu önceden bilinmediği için gelen her yeni girdi olası bir bitiş noktası olarak varsayılır. Daha sonra sınırları önceden tanımlanmış bir aralık taranarak olası başlangıç noktaları aranır. Bu aralıktan başlayan tüm diziler için olabilirlikler hesaplanır ve eğer bir dizi için HMM modellerinden birinin verdiği olabilirlik eşik modelin verdiği olabilirliği aşarsa söz konusu hareket tanınmış olarak kabul edilir.



Şekil 5 El Hareketleri

3. Uygulama

Yaygın kişisel bilgisayar uygulamaları fare tıklamaları ve klavye girdileri ile yönetilen basit arayüzlere sahiptir. Oyun ve çizim programları gibi daha gelişmiş uygulamalar farenin hareketini de göz önünde bulundurlar. Burada önerilen sistem bu tür gelişmiş uygulamalara arayüz oluşturabilmek için tasarlanmıştır. Bu çalışmada gerekli verileri alabilmek amacıyla bu arayüzü ticari bir çizim programı ile bağlantılandırdık.

Söz konusu çizim programı, komutları fare ile uygulama düğmeleri seçtiğinde almakta ve bu komutları program tuvali tıklandığında uygulamaktadır. Geliştirdiğimiz arayüzde, hareket tanıma, komutların seçilmesinde, el izleme ise bu komutların uygulanmasında kullanılmaktadır. Fare tıklaması, Şekil 2’de görüldüğü gibi kullanıcı elini kapatıp baş parmağını açtığı anda algılanmaktadır. Fare imleci, bulunan üç boyutlu el koordinatlarının ekran üzerindeki iki boyutlu izdüşümleri alınarak yönetilir ve bir el hareketi tanındığında o hareketle bağlantılı olan komut işleme konur.

Geliştirdiğimiz arayüz, kullanıcıya belirteci değiştirme, kamera kalibrasyon matrislerini bulma, sistem parametrelerini değiştirme ve yeni hareketleri öğrenme gibi işlemlerin yapılabilmesi için seçenekler sunmaktadır. Bu sayede el hareketleri kullanıcı tarafından program içinden eğitilebilir ve komutlara bir konfigürasyon dosyası yardımıyla bağlanabilir.

4. Sistem Başarısı

Sisteme Şekil 5’te görülen sekiz el hareketini tanıttık ve bunları Tablo 1’de görüldüğü gibi çizim programının komutları ile bağlantılandırdık. Şekil

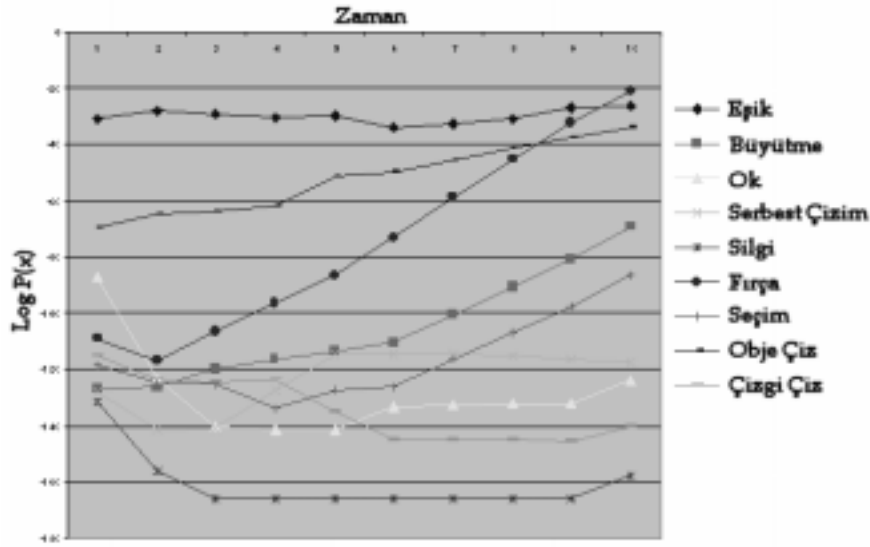
6, zaman içinde değişik hareketlere karşılık gelen HMM’lerin olabirliklerinin değişimini vermektedir. Herhangi bir HMM’in olabirliği uyarlanır eşik modelinin değerini geçtiğinde, o hareket tanınmış olur.

Sistemi test ederken, çok kısa dizilerin ve birbirinden farklı en az üç yönsel sembol içermeyen dizilerin hatalı bir şekilde eşik modelinden daha yüksek olabirlikler verdiğini ve bunun yanlış tanıma sebepleri olduğunu gözlemledik. Bu nedenle sistemimizde bu tür diziler sınıflandırma işlemine katılmamaktadır. Gözlemlediğimiz başka önemli bir nokta ise tanımların %98’inin seçilen aralıktaki tanıma şartlarına uyan en kısa dizi için gerçekleşmesidir. Bu, sistemin performansının uzun dizilerin gözardı edilmesinden etkilenmediğini göstermektedir.

Sistemin 160 deney sonucunda gösterdiği hareket tanıma performansı Tablo 1’de görüldüğü gibi toplam 2 yanlış tanıma ile %98.75’tir.

5. Sonuç

Bu çalışmada üç boyutlu el hareketlerinin gerçek zamanlı takibi ve tanınması üzerinde çalıştık. Öncelikle belirteç çeşitli görüntü işleme algoritmaları kullanılarak kameralardan gelen görüntülerde ayırdedildi. Ardından her iki kameradan gelen görüntüler kullanılarak üç boyutlu geriçatma işlemi uygulandı. Daha sonra elde edilen gezinge önceden eğitilmiş HMM’lere gönderildi ve olabirliği eşik modelini geçen hareketler tanındı. Gerçek zamanlı olarak %98.75 tanıma oranı ile çalışan bu sistem etkileşimli bir hareketsetel arayüz örneği olarak önerilmektedir.



Şekil 6 Modellerin Olabilirlikleri

Hareket	Durum	Öğrenme	Deneme	Doğru	Yanlış	Yüzde	Araç
	8	65	20	20	0	100	Büyütme
	4	56	20	20	0	100	Ok
	8	59	20	20	0	100	Serbest Çizim
	4	59	20	20	0	100	Silgi
	3	60	20	20	0	100	Fırça
	4	56	20	19	1	95	Seçim
	8	57	20	19	1	95	Obje Çiz
	3	57	20	20	0	100	Çizgi Çiz

Tablo 1 Hareketlerin Tanınma Oranları

KAYNAKLAR

- [1] L.R Rabiner, "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proc IEEE*, vol.77, pp.257-285, 1989
- [2] H. Lee and J.H Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 961-973, 1999.
- [3] T. Starner and A. Pentland, "Real Time American Sign Language Recognition from Video

Using Hidden Markov Models," Technical Report TR-375, MIT's Media Lab., 1995

[4] K. Oka, Y.Sato and H. Koike, "Real-Time Fingertip Tracking and Gesture Recognition", *IEEE Computer Graphics and Applications*, November-December, 2002, pp. 64-71.

[5] Y. Sato, M. Saito, and H. Koike, "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, March 2001.