

BELİRLİ SÜRE MODELLERİ İLE İZOLE EL HAREKETİ TANIMA EXPLICIT DURATION MODELS FOR ISOLATED HAND GESTURE RECOGNITION

Cem Keskin¹, Ali Taylan Cemgil¹, Lale Akarun¹

¹ Bilgisayar Mühendisliği Bölümü, Boğaziçi Üniversitesi, Türkiye

keskinc@cmpe.boun.edu.tr

taylan.cemgil, akarun@boun.edu.tr

Özetçe

Bu bildiri de gerçek zamanlı izole el hareketi tanıma problemi için belirli süre modellerinin (BSM) başarımları ve hızı incelenmiştir. Öncelikle el hareketlerinin karakteristik özelliklerinin ayrıntılı bir analizi ile homojen saklı Markov modellerinin (SMM) yetersizlikleri irdelenmektedir. Özellikle SMM'lerde durum sürelerinin açık modellenmesinin gerekliliğine işaret edilip, bu probleme çözüm olarak BSM önerilmektedir. Bu önermeyi sınamak için BSM tabanlı bir el hareketi tanıma çatısı geliştirilmiş, 10 adet 3B el hareketi içeren bir veri kümesi oluşturulmuş ve önerilen yöntemle beraber SMM'lerin, saklı koşullu rasgele alanların (SKRA) ve girdi çıktı SMM'lerin (GÇSMM) izole el hareketi dizileri üzerindeki hareket sınıflandırma başarımları 5x2 çapraz geçerlilik sınaması yöntemi ile test edilmiştir. Tanıma başarımlarını maksimize eden parametreler için modellerin test süreleri karşılaştırılmış ve gerçek zamanlı sistemler için en uygun modelin BSM olduğu sonucuna varılmıştır.

Abstract

In this paper we test the recognition efficiency of explicit duration models (EDM) for isolated gesture recognition. First, through a careful analysis of the characteristics of hand gesture patterns, the shortcomings of homogeneous hidden Markov models (HMM) are pointed out. Next, EDM is proposed as an efficient method to model durations. Finally, to validate these claims, an EDM based framework is developed and tested along with HMMs, hidden conditional random fields and input-output HMMs, on a database consisting of 10 3D hand gestures using 5x2 cross validation. By comparing the testing times of models using parameters that maximize the recognition rates, it is concluded that EDMs are better suited for real-time applications than the other models.

1. Giriş

El hareketi tanıma problemi doğal insan bilgisayar etkileşiminde önemli bir role sahiptir. Ancak görü tabanlı el hareketi tanıma, sensörlerden gelen bilgilerin gürlüğü seviyesi yüzünden özellikle zor bir problemdir. Işıklandırma şartlarının görüye etkisi, elin her çerçevede verimli bir şekilde

ayrıştırılması ve üçüncü boyut bilgisine ulaşmak için birden fazla kamera ve belirli bir kamera düzeni gerekmesi bu yöntemin yaygınlığını azaltmıştır. Yakın tarihte ise kızılötesi lazerler yöntemiyle derinlik algılayabilen sensörlerin kullanımının sıklaşması üzerine el hareketlerinin özellikle oyun tarzı uygulamalarda kullanımı hızla yaygınlaşmaya başlamıştır.

İnsan bilgisayar etkileşimindeki el hareketleri, önceden tanımlanmış bir grup hareketin önceden belirlenmemiş bir sırayla uygulanmasıyla oluşturulur. El hareketi tanıma sistemlerinin bu hareketleri gerçek zamanlı olarak yakalaması ve tanınması gerekir. Bunun için sisteme belirlenmiş bir grup el hareketinin önceden tanıtılmış olması gerekmektedir.

El hareketi tanıma sistemleri, ya tanınmış hareketlere denk gelen sinyallerin sürekli dizilerden önceden ayıklanmış ve izole edilmiş olduğunu varsayar ve ayrı olarak tanır, ya da bu hareketleri sürekli diziler içerisinde tanımaya çalışır. İzole el hareketi tanıma sistemleri el hareketlerinin ne zaman başlayıp ne zaman bittiğinin belli olduğunu varsayar. Ancak her uygulamada el hareketlerinin başlangıç ve bitiş noktaları belirlenmiş değildir ve tanıma esnasında sistemin tanınan hareketlerin yerini de saptaması gerekir. Bunun yanı sıra, istem dışı veya iki hareketin bağlanması amaçlı yapılan hareketlerin ayıklanıp gözardı edilmesi gerekmektedir. Sürekli el tanıma problemi, işaret dili tanıma gibi uygulamalar için özellikle önemlidir.

El hareketleri, elin gezinmesinin ve şeklinin spesifik birer kombinasyonundan oluşur. Dolayısıyla sürekli el hareket dizileri, birbirine bağlı belirli örüntü veya sinyal dizileri olarak düşünülebilir. El hareketlerinin genelde Markov özelliğine sahip kısmen gözlemlenebilir stokastik süreçlerin çıktısı olduğu varsayılır [1]. Markov özelliği, anlık gözlemlerin olasılık dağılımlarının sadece sistemin o anki durumuna koşullandırılmış olması demektir. Zaman serisi modelleyebilen Markov zincirlerinin en basiti ve en yaygını SMM'dir. El hareketi tanıma problemi için yapılan çalışmaların büyük çoğunluğu da SMM ve türevleri üzerinde yoğunlaşmıştır. Belirli SMM türevlerinin standart SMM'lere göre başarımları bir çok çalışma tarafından ortaya konmuş, ancak alternatif yöntemlerin SMM'lere göre daha başarılı olmasının sebebi ayrıntılı şekilde analiz edilmemiştir. Ayrıca başarımları yüksek bir model, eğer karmaşıklığı yüksekse gerçek zamanlı el tanıma sistemleri için uygun olmayabilir. Bu çalışmada el hareketlerini derinlemesine

analiz ederek izole ve sürekli el hareketi tanıma problemleri için nasıl modellerin daha uygun olacağı çıkarımında bulunuyoruz ve hızlarını karşılaştırıyoruz.

1.1. El hareketleri için genel bir formülasyon

Farklı kişiler aynı el hareketini birbirinden oldukça farklı şekilde yaptıkları halde, hareketi izleyenler bu hareketleri algılamada bir sorun yaşamazlar. Dolayısıyla, gözlemlenen el niteliklerinin kesin reel sayısal değerlerinin önemi yoktur, ve birbirine benzer gözlemlerin nicemleme yoluyla tek bir sembole indirgenmesinin tanıma başarımında önemli bir etkisi olması beklenmez. Öte yandan bu nicemleme işlemi genel el hareketlerin analizini kolaylaştırması açısından anlamlıdır.

Görüntü tabanlı el hareketi tanıma sistemlerinde elin hareket ve şekil bilgisi kameralar ve izleme algoritmaları yoluyla tüm çerçevelerde bulunur ve her imge için bir öznitelik vektörü oluşturulur. El hareketlerini modelleme problemi, bu öznitelik vektörlerinin hareket sınıflarına koşullandırılmış olasılık dağılımlarını öğrenmeye eşdeğerdir. Bu özniteliklerin nicemlenmesi yolu ile el hareketi sinyallerini birer sembol ya da kod sözcüğü dizisi olarak göstermek olasıdır. İzole edilmiş bir el hareketini şu şekilde göstermek mümkündür:

$$t_1^{e_1 m} t_2^{e_2 m} t_3^{e_3 m} \dots t_k^{e_k m} \quad (1)$$

Burada her t_i belirli bir aralıktaki sembolü tanımlar. Komşu olmadıkları sürece t_i ve t_j aynı sembol olabilir. Her t_i sembolünün üssü olan $e_i m$ sayısı o sembolün kaç kere tekrarlandığını göstermektedir. e_i değeri karakteristik üstür ve el hareketinin dinamiklerine bağlıdır. m değeri ise uygulanan el hareketinin hıza, hareketin büyüklüğüne ve sensörlerin çerçeve hızına bağlıdır ve tüm semboller tarafından paylaşılır. Ancak bu formülasyonun geçerli olabilmesi için elin hareket esnasında izleyebileceği alternatif gezinmeler ve sahip olabileceği farklı el şekilleri olmayacağı varsaymak gerekmektedir. Bu varsayım, kullanılan modelin sol-sağ mimariye sahip olmasına olanak tanıdığı için önemlidir. Çoğu uygulamada bu doğru bir varsayımdır, ama tersi örnekler de mevcuttur. Örneğin havaya çizilen harf veya rakamlar gibi semantik anlamı olan hareketler farklı insanlar tarafından alışkanlıkları doğrultusunda farklı şekillerde uygulanabilir.

Sürekli el hareketleri, Denklem 1'deki dizilerin arka arkaya gelmesiyle oluşur. Ancak her el hareketi farklı hızda ve boyutta uygulanabileceği için m değeri farklı el hareketleri arasında paylaşılmaz. Ayrıca her iki el hareketi arasında tanımayan veya istem dışı yapılmış el hareketleri bulunabilir. El hareketlerinin baş ve sonlarının da birbirlerine *ulanması* mümkündür.

1.2. El hareketlerinin modellenmesi

1.2.1. Saklı Markov modelleri

Denklem 1'deki sembol dizilerini modelleyen bir SMM'in her saklı durumu bu sembol dizilerinin belirli bir kısmından sorumludur. Örneğin SMM'in k tane saklı durumu varsa, ideal bir eğitim sonunda her durumun bir $t_i^{e_i m}$ bölütüne denk gelmesi, ve sistem bu durumlardayken t_i sembolünü üretme ihtimalinin diğer sembollerinkinden yeterince büyük olması beklenir. $e_i m$ sürelerini ise durumlarda ne kadar zaman geçirileceği, dolayısıyla başka durumlara geçiş olasılıkları belirler. Bu

olasılıkların değişmediği homojen SMM'lerde bütün süreler birer geometrik dağılım ile modellenir. Bu süreler her durum için birbirinden bağımsızdır; dolayısıyla ortak üs olan m değerini naif Bayes varsayımı yapan homojen SMM'lerle modellemek mümkün değildir. Ancak hareketlerin hız ve boyutlarının tek doruklu bir olasılık dağılımından gelmesi şartıyla bu sorun ortadan kaldırılabilir. Bu varsayımın özellikle birden fazla kişi tarafından kullanılacak sistemler için geçerli olması beklenmez.

1.2.2. Saklı koşullu rasgele alanlar

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. SMM'lerin ayırıcı karşılıkları koşullu rasgele alanlardır (KRA) [2]. KRA'lar naif Bayes varsayımında bulunmaz; her durum örtüşen bir grup gözlemden çıkarılmış özniteliklere koşullandırılmıştır. Ancak KRA'lar bir sınıfın iç dinamiklerini değil, sınıflar arası dinamikleri modeller ve dolayısıyla zaman serisi sınıflandırmaya uygun değildir. Bu nedenle saklı KRA'lar (SKRA) [3] ve gizli dinamik KRA'lar (GDKRA) [4] 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.

SKRA'lar durum sürelerini öğrenmek yerine gözlemleri durum geçişleriyle ilişkilendirir. Bu özellik pozitif örneklerin tanınma başarımını artırır, ancak yanlış kabul oranının da yükselmesini sağlar. Eğer veri kümesindeki farklı sınıfların örnekleri birbirinden yeterince ayrıksa SKRA yüksek başarımları ulaştırır.

1.2.3. Girdi-çıkıtlı saklı Markov modelleri

Girdi-çıkıtlı SMM'ler (GÇSMM) üretici ve ayırıcı model melezleri olarak izole el hareketi tanımda yüksek başarımlar göstermiştir [5]. Bu modeller de SKRA'lar gibi durum geçiş olasılık dağılımlarını gözlemlerin bir fonksiyonundan oluşan bir girdi dizisine koşullandırır [6]. Bunun yanı sıra gözlemlerin koşullu olasılık dağılımları da aynı veya başka bir girdi dizisine koşullandırılır. GÇSMM'ler özde homojen olmayan SMM'lerdir. Gözlem ve durum geçiş olasılıkları her çerçevede lokal modeller kullanılarak girdi dizisinden hesaplanır. Lokal model olarak her türlü fonksiyonu ya da modeli kullanmak mümkündür. GÇSMM diğer modellerden farklı olarak bir uzman tarafından tasarlanmalı, hangi gözlemlerin girdi, hangilerinin çıktı olacağına ve ne karmaşıklıkta lokal modeller kullanılacağına karar verilmelidir. Lokal model olarak örneğin radyal taban fonksiyonları veya çok katmanlı perseptronlar kullanılabilir. GÇSMM'ler SMM'lere göre daha karmaşıktır ve eğitimleri daha çok örnek gerektirir.

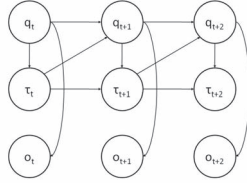
1.2.4. Saklı yarı Markov modelleri

El hareketi dizilerini tanımaya daha yatkın bir modelin hem süreleri açık şekilde modellemesi, hem de bu sürelerin birbirine bağlı olmasına olanak vermesi beklenir. Dolayısıyla saklı yarı Markov modeller (SYMM) el hareketi tanıma problemi için uygun bir alternatiftir. SYMM'ler SMM'lere benzemelerine

karşın, SYMM saklı durumları tek bir gözlem üretmek yerine boyları belirli olasılık dağılımlarından gelen birer gözlem dizisi üretir [7, 8]. SYMM'ler SMM'lere benzer bir grafik modeline sahip üretici modellerdir. Ancak farklı olarak SYMM durumları tek bir sembol yerine bir sembol dizisi oluşturur. Değişik SYMM'ler bu dizileri farklı şekilde üretebilir. Örneğin bölüt modellerinde (*segment model*) her durum ayrı birer SMM gibi davranır. Belirli süreli modeller (BSM) (*explicit duration model*) ise her saklı durumda tam olarak ne kadar kalacağını eklenti bir sayaç durumu ile belirler. Bu çalışmada el hareketlerini oluşturan her parçanın süresinin belirli şekilde modellenmesinin tercih edilmesinden dolayı BSM'ler üzerinde durulmuştur.

BSM'lerde o_t gözlemleri SMM'lerdeki gibi s_i değerlerinden birini alan q_t saklı durumlarına koşullandırılmıştır. Ancak BSM'lerde durum geçişleri tam sayı değer alan ve kalan süreyi belirten τ_t değerine bağlıdır. τ_t bir sayaç görevi görerek her adımda deterministik şekilde azalır. Durum geçişlerine ancak τ_t 'nin değeri sıfır olduğunda izin verilir. Sistem s_m durumuna geçiş yaptığında τ_t , $p_m(d)$ olasılık dağılımından çekilmiş d değerini alır. Bu sürelerin olasılık dağılımları parametrik şekilde modellenebileceği gibi, parametrik olmayan bir dağılım kestirmek de mümkündür. Geri kalan parametreler SMM ile ortaktır. s_m durumunda iken v_k sembolünü gözleme ihtimali $b_m(v_k)$ 'dir. τ_t sıfırlandığında s_m durumundan s_n durumuna geçme ihtimali a_{mn} 'dir. Sistemin s_m durumundan başlama ihtimali de π_m 'dir. BSM grafik modeli Şekil 1'de gösterilmiştir.

BSM'ler sürelerin belirli şekilde modellenmesi problemini çözer, ancak süreler hız ve boyuta koşullandırılmadığı sürece halen birbirlerinden bağımsızdır.



Şekil 1: Belirli süre modeli

2. Yöntem

k tane saklı duruma sahip bir BSM ile Denklem 1'deki izole el hareketi modellendiği takdirde, her s_i durumunun bir bölütü, her $p_i(d)$ olasılık dağılımının da o bölütün süresini modellemesi beklenir. Sabit gezinge varsayımı doğru ise, bu dizileri bir sol-sağ BSM ile modellemek mümkündür.

BSM tabanlı bir el hareketi tanıma çatısı, her hareket sınıfının ayrı bir BSM ile modellenmesi yöntemiyle tasarlanabilir. Bu yöntem, SMM ile sınıflandırma yapma ile aynıdır. Başı ve sonu belli, sınıflandırılmamış bir el hareket dizisi geldiğinde sistemde tanımlı tüm BSM'ler tarafından değerlendirilerek sınıf olabirlikleri hesaplanır ve en yüksek değeri veren BSM'nin sınıfı etiket olarak seçilir. İzole el hareketi tanımadaki çoğu grafik model aynı şekilde kullanılabilir. Her hareket sınıfında yeterli sayıda eğitim örneği alınır ve modeller ayrı ayrı bu pozitif örneklerle eğitilir.

BSM'ler beklenti maksimizasyonu yöntemi ile eğitilebilir.

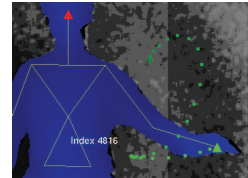
Tüm τ_t değerleri ile s_i değerlerinin kombinasyonuna bir SMM durumu atamak mümkün, ancak verimsizdir çünkü bu şekilde oluşturulan bir SMM'in durum geçiş matrisi seyrek-tir. Yu tarafından bu seyrekliği kullanarak verimli eğitim ve değerlendirme yöntemi önerilmiştir [7]. Bu yöntem SMM eğitimde kullanılan Baum-Welch algoritmasının belirli sürelerle uyarlanmış halidir, ve bu çalışmanın deneyleri için bu yöntem kullanılmıştır.

3. Deneyler

BSM'lerin başarımlarını başka modellerle karşılaştırmak ve tanıma hızını ölçmek için zorlu bir veri kümesi topladık ve SMM, GÇSMM ve SKRA ile de test ettik. Bu bölümde veri kümesinin özellikleri ve deney yöntemimiz açıklanmakta, son olarak da sonuçlar verilmektedir. GÇSMM'lerin başarımlarını genelleştirebilmek için alabildiğine genel bir mimari kullanılmış ve girdi olarak sadece zaman bilgisi verilmiştir. Bu haliyle GÇSMM homojen olmayan ve naif Bayes varsayımı içermeyen bir SMM gibi davranır.

3.1. Veri kümesi

Tanıma başarımının görü modüllerinden bağımsız olmasını sağlamak için kızılötesi derinlik sensörüne sahip Kinect kamerası kullanıldı. 12 farklı kişi tarafından sıfırdan dokuza kadar olan rakamlar 2B bir düzlem üzerinde onar kere uygulandı ve her sınıftan 120 örnek toplandı. Rakamların uygulanması konusunda kişiler belirli bir gezingeyi takip etmeye zorlanmadı. Dolayısıyla sabit gezinge varsayımının geçerli olmadığı, gezinmelerin varyansının oldukça yüksek olduğu zorlu bir veri kümesi yaratıldı.



Şekil 2: Bu şekilde iki (2) rakamının bir örneğinin uygulanma anının Kinect kamerasının derinlik sensörleri tarafından yakalanma anı görüntülenmektedir.

3.2. Öznitelik seçimi

Sabit gezinge varsayımına uyan veri kümeleriyle çalışıldığında öznitelik olarak elin iki çerçeve arasında yaptığı hareketin açısı ve boyu nicemlenerek kullanılabilir [9]. Ancak bu veri kümesinde elin hareketinden çok açığa çıkan son şeklin önemli olduğu açıktır. Dolayısıyla bağıl hareketler yerine normalize edilmiş pozisyon bilgisinin başarımlarını arttıracığı öngörülebilir. Bu çalışmada iki tip öznitelik de eğitimde kullanılmış ve sonuçları karşılaştırılmıştır.

3.3. Eğitim

Deneylerde kullanılan SMM'ler standart Baum-Welch algoritmasıyla, BSM'ler Baum-Welch'in belirli sürelerle geliştirilmiş türeviyle [7], GÇSMM'ler ve lokal modelleri olan çok katmanlı perseptronlar geliştirilmiş beklenti

maksimizasyonu [5] yöntemiyle, SKRA'lar ise Broyden-Fletcher-Goldfarb-Shanno [3] yöntemi ile eğitilmiştir. SMM'lerin ve BSM'lerin eğitim süresine göre GÇSMM ve SKRA daha uzun sürmektedir. Arada yaklaşık iki büyüklük derecesi fark vardır.

3.4. Model parametre optimizasyonu

Deneysel 5x2 çapraz geçerlilik yöntemi ile uygulanmış, her aşamada sınıflardan rasgele seçilmiş 60 örnek ile eğitim, geri kalan 60 tane ile de test yapılmıştır. Model seçme yöntemi olarak her model parametresi belirli aralıklarda değiştirilerek test sonuçları en yüksek olanların parametreleri seçilmiştir. SMM'ler ve BSM'ler için tek parametre saklı durum sayısı N 'dir. SKRA'ların parametreleri saklı durum sayısı N ve her durumun koşullandırıldığı gözlem sayısını belirten pencere genişliği w 'dir. GÇSMM'ler için ise [5] çalışmasında önerildiği gibi lokal model olarak çok katmanlı perseptron kullanılmış, girdi dizisi olarak da sadece normalize edilmiş zaman bilgisi kullanılmıştır. GÇSMM parametreleri, saklı durum sayısı N ve perseptronların gizli nöron sayısı H 'dir. Ayrıca her model iki farklı öznitelik kümesiyle test edilmiştir. Tabloda bağlı hareket öznitelikleri F_H , normalize edilmiş pozisyon öznitelikleri de F_P sembolü ile gösterilmiştir. Durum sayıları 1 ve 20 arasında, gizli nöron sayıları 5 ve 10 arasında, pencere genişliği ise 0 ile 4 arasında test edilmiştir. Optimum parametreler Tablo 1'de listelenmiştir. Bütün modellerin yüksek sayıda durum sayısına ihtiyaç duymuş olmaları veri kümesinin zorluluğunu doğrulamaktadır.

	F_H		F_P	
SMM	N= 16		N= 16	
BSM	N= 15		N= 19	
GÇSMM	N= 9	H= 10	N= 8	H= 5
SKRA	N= 9	w= 3	N= 7	w= 3

Tablo 1: Optimum model parametreleri. Bu tabloda N durum sayısı, H gizli nöron sayısı ve w pencere genişliğini tanımlar.

3.5. Sonuçlar

Optimum parametrelerle yapılan 5x2 çapraz geçerlilik testlerinin sonuçları Tablo 2'de verilmiştir. Hareket bilgisi içeren F_H özniteliklerine göre pozisyon bilgisi içeren F_P özniteliklerinin daha ayırıcı olduğu sonuçlarca doğrulanmıştır. BSM'nin başarımları beklendiği üzere SMM'e göre daha yüksektir, ancak GÇSMM ve SKRA bu veri kümesinde daha yüksek başarımlara sahiptir. Veri kümesindeki varyansın fazlalığı süreleri modellemeye çalışan SMM ve BSM'nin başarımlarını düşürürken durum geçiş anlarını gözlemlere koşullandıran SKRA varyanstan diğer modeller kadar etkilenmemiştir. Ancak veri kümesindeki sınıf sayısı yeterince çok olmadığı için SKRA'lardan beklenen yanlış kabul oranı tam olarak sonuçlar tarafından yansıtılmamıştır. GÇSMM ise sadece normalize edilmiş zamana koşullandırıldığı halde, yani gözlemlerdeki değişiklik bilgisini kullanmadığı halde SKRA ile yaklaşık aynı performans göstermiştir.

Tablo 2'de 60 örneği test etmek için geçen süreler saniye cinsinden listelenmiştir. Bu değerlere göre, GÇSMM ve SKRA daha yüksek başarımlarına rağmen SMM ve BSM'ye göre

en az bir mertebe daha yavaştır. Gerçek zamanlı çalışması gereken bir sistem için SMM'lerden daha yüksek başarımla orana sahip, hız kaybı da göz ardı edilebilir olan BSM'lerin daha verimli olduğu iddiası sonuçlarca desteklenmektedir.

	F_H	F_P	Süre
SMM	73.5%	89.7%	1.04s
BSM	76.83%	91.17%	1.11s
GÇSMM	77.50%	94.33%	42.5s
SKRA	79.3%	95.17%	13.38s

Tablo 2: Veri kümesi üzerinde tanıma başarımları ve test süreleri.

4. Vargı

Bu çalışmada dört farklı grafik modelini izole el hareketi tanıma problemi için başarımları ve test hızları açısından test ettik. Yapılan deneyler sonucunda SKRA ve GÇSMM'lerin başarımlarının BSM ve SMM'lere göre daha yüksek olduğu, ancak gerçek zamanlı sistemler için yavaş kaldığı bulgusuna vardık. Gerçek zamanlı sistemler için SMM'lerden daha yüksek başarımları gösteren ve hızı yaklaşık aynı olan BSM'ler önerilmektedir.

5. Kaynakça

- [1] Vladimir Pavlovic, Rajeev Sharma, and Thomas S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Tran. on Patt. Anal. and Machine Intel.*, vol. 19, no. 7, pp. 677–695, 1997.
- [2] J. D. Lafferty, A. McCallum, and 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*. 2001, pp. 282–289, Morgan Kaufmann Publishers Inc.
- [3] Sy Bor Wang, Ariadna Quattoni, Louis-Philippe Morency, and David Demirdjian, "Hidden conditional random fields for gesture recognition," in *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, DC, USA, 2006, pp. 1521–1527, IEEE Computer Society.
- [4] Louis-Philippe Morency, Ariadna Quattoni, and Trevor Darrell, "Latent-dynamic discriminative models for continuous gesture recognition," *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, vol. 0, pp. 1–8, 2007.
- [5] C. Keskin and L. Akarun, "Stars: Sign tracking and recognition system using input-output hmms," *Pattern Recogn. Lett.*, vol. 30, pp. 1086–1095, September 2009.
- [6] Y. Bengio and P. Frasconi, "Input-output HMM's for sequence processing," *IEEE Transactions on Neural Networks*, vol. 7, no. 5, pp. 1231–1249, September 1996.
- [7] Shun-Zheng Yu and Hisashi Kobayashi, "Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden markov model," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1947–1951, 2006.
- [8] Kevin P. Murphy, "Hidden semi-markov models," Tech. Rep., 2002.
- [9] C. Keskin, A. Erkan, and L. Akarun, "Real time hand tracking and 3d gesture recognition for interactive interfaces using hmm," in *In Proceedings of International Conference on Artificial Neural Networks*, 2003.