

Saklı Markov Model Karışımları için Spektral Öğrenme

Spectral Learning of Mixtures of Hidden Markov Models

Yusuf Cem Sübakan¹, Oya Çeliktutan¹, Ali Taylan Cemgil², Bülent Sankur¹
Elektrik-Elektronik Mühendisliği¹, Bilgisayar Mühendisliği²

Boğaziçi Üniversitesi, Bebek 34342

{cem.subakan; oya.celiktutan; taylan.cemgil; bulent.sankur}@boun.edu.tr

Özetçe —Bu çalışmada, Saklı Markov Modeli (SMM) olarak modellenen zaman serilerinin topaklandırılması için yeni bir yöntem önerilmektedir. Topaklardaki SMM parametrelerinin güncellenmesi için son yıllarda yapay öğrenme literatüründe popüler olmaya başlayan saklı değişken modelleri için spektral öğrenme yöntemleri kullanılmaktadır. Spektral yöntemler, alışlagelmiş beklenti-enbüyütme yaklaşımının aksine, saklı değişken modellerinde tek adımda parametre kestirimi yapmamızı sağlar. Bu sebeple, önerdiğimiz yöntem hesap karmaşıklığı bakımından alışlagelmiş yöntemlerle SMM-topaklandırmaya göre hesap karmaşıklığı bakımından daha ucuzdur.

Anahtar Kelimeler—Saklı Markov Modeli, Karışım Modeli, Spektral Öğrenme

Özet—In this work, we propose a novel approach for clustering Hidden Markov Models (HMMs). We use spectral learning for latent variable models to learn HMM parameters in each cluster. Unlike conventional expectation-maximization algorithms, spectral learning enables us to do parameter estimation in latent variable models without iterating, in local optima free fashion. For this reason, our algorithm is computationally cheaper than clustering HMMs with conventional approaches such as EM.

Keywords—Hidden Markov Model, Mixture Model, Spectral Learning

I. GİRİŞ

Saklı Markov Modeli (SMM), zaman serisi modellemek için otomatik öğrenme literatüründe yaygın olarak kullanılan bir yaklaşımdır. Bu modelde, gözlemlerin saklı bir Markov durum dizisine koşullu bir olasılıksal kurala bağlı olarak üretildikleri varsayılır. SMM'lerin geniş bir uygulama alanı olmakla beraber, bunlara örnek olarak konuşma tanıma [1], müzik analizi [2], zaman serisi topaklandırma [3], [4], [5], [6] verilebilir. Bu çalışma SMM olarak modellenen zaman serilerinin topaklandırılması ile ilgilidir. Bu amaçla, son dönemde popüler olmaya başlayan saklı değişken modelleri için spektral öğrenme yöntemlerinden [7], [8], [9] yararlanılmaktadır.

Saklı Markov Modeli topaklandırma konusunda literatürde çok fazla çalışma bulunmamaktadır. Çalışmaların çoğu, olabilirlik enbüyütmeye dayalı parametrik topaklandırma yapmaktadır [3], [5], [4]. Zaman serileri arasındaki benzerliği bir çizge

modeline oturtan ve spektral topaklandırmayı [10] kullanarak topaklandırma yapan bir yöntem de mevcuttur [6]. Bizim çalışmamızda, [6]'daki gibi bir çizgesel spektrum yöntemi değil, gözlemlenebilir momentlerin özdeğer spektrumuna dayalı bir yöntem kullanılmıştır.

Literatürdeki parametrik SMM karışımı öğrenme algoritmalarının birçoğunda parametre kestirimi için beklenti-enbüyütme algoritması (EM) kullanılmaktadır. Beklenti-enbüyütme algoritması saklı değişken modelleri için en sık kullanılan parametre kestirimi algoritmalarından biri olmakla beraber, yerel maksimumlara takılma eğilimi göstermekte ve döngü yapmayı gerektirmektedir.

Beklenti enbüyütme kullanarak SMM karışımı öğrenmeyi ele alalım: Her döngüde bütün topaklar için beklenti-enbüyütme algoritması ile parametre kestirimi yapmamız gereklidir. Bizim önerimiz, beklenti-enbüyütme algoritması yerine tek adımda spektral öğrenme kullanarak parametre kestirimi yapmaktır. Bu yaklaşımın başlıca getirileri hesap yükünün azaltılmasıdır ve parametre kestirimi için ilkendirme yapmak zorunda kalmamaktır.

Saklı değişken modelleri için spektral yöntemler, verinin düşük dereceli momentlerini kullanarak model parametrelerinin kestirilmesini özdeğer-özvektör ayrıştırma problemine indirgeyebilir. Bu yaklaşım istatistikteki moment eşleme yöntemi ile ilişkilidir. Bu alandaki ilk göze batan çalışma Saklı Markov Modellerinde olabilirlik kestirimi için [7] yapılmıştır. Ardından karışım modelleri, LDA, ICA gibi diğer popüler modellere de uygulanmıştır [8], [9]. Bu yöntemlerin genel saklı ağaç yapılarına genellenmesi de mevcuttur [11].

Önerilen yöntem, örnek uygulama olarak insan bedeninin devinimlerini yakalama (MoCap: Motion Capture) verisinin modellenmesi için, bir başka deyişle insan vücudu üstünde alınan nirengi noktalarının uzay-zamandaki değişimlerinin modellenerek insan hareketlerinin betimlenmesi ve topaklandırılması amacıyla kullanılmıştır. Önerdiğimiz yöntem, bu verinin beklenti en-büyütme yaklaşımıyla karşılaştırılabilir performans vermekte ve süre açısından 3-4 kat kazanç sağlamaktadır.

Bildirinin akışı şu şekildedir. 2. Bölümde SMM ve SMM için spektral öğrenme yöntemi tanımlanmıştır. 3. Bölümde ise SMM karışımları ve SMM karışımlarını öğrenme yöntemleri

verilmiştir. 4. Bölüm karşılaştırmalı deneysel sonuçları ve 5. Bölüm ise gelecek çalışmayı özetlemektedir.

II. TEMEL TANIMLAR

A. Simgelem

Makalede başlıca kullanılan notasyon aşağıdaki gibidir:

- Değişken listeme: $x_{1:T} = \{x_1, x_2, \dots, x_T\}$.
- N boyutlu olasılık simpleksi: $\Delta^{N-1} = \{(p_1, p_2, \dots, p_N) \in \mathbb{R}^N \mid p_n \geq 0 \forall n, \sum_n p_n = 1\}$.
- Dış çarpım: $(x \otimes y)_{ij} = x_i y_j$.
- N boyutlu standart taban vektör kümesi: $\mathcal{E}^N = \{e_1, e_2, \dots, e_N\}$, e_n n 'inci elemanı 1, geri kalanı 0 olan bir vektördür.
- Gösterge fonksiyonu: $\forall (x, y) \in \mathbb{Z}$,
 $[x = y] = 1$, $[x \neq y] = 0$.

B. Saklı Markov Modelleri

Bir SMM'ye göre, gözlem dizisi $x_{1:T}$ aşağıdaki gibi bir olasılık dağılımdan gelir:

$$p(x_{1:T}|\theta) = \sum_{h_{1:T}} \prod_{t=1}^T p(x_t | h_t) p(h_t | h_{t-1}) \quad (1)$$

$h_{1:T}$ saklı durum dizisi ve $\theta = (A, O, \pi)$ model parametreleridir ve aşağıdaki gibi tanımlanmıştır:

$$\pi_u = p(h_1 = u | h_0) = p(h_1) \quad (2)$$

$$A_{uv} = p(h_t = u | h_{t-1} = v), \quad t > 1 \quad (3)$$

$$O_u = \mathbb{E}[x_t | h_t = u] \quad (4)$$

Bu tanımlarda, ilk durum dağılımı $\pi \in \Delta^{K-1}$, durum geçiş matrisi $A \in \mathbb{R}^{K \times K}$ ve $A_v \in \Delta^{K-1}$, gözlem matrisi $O \in \mathbb{R}^{D \times K}$. K saklı durum sayısı, D ise gözlem vektörünün boyudur. Gözlem modeli $p(x_t | h_t)$ 'nin ne olacağı belirtilmemiştir (sürekli, ayrık hatta melez bile olabilir), uygulamaya uygun olarak seçilmeye bırakılmıştır.

C. Saklı Markov Modelleri için Spektral Öğrenme

Amaçlanan, gözlemlenen $x_{1:T}$ dizisinin düşük dereceli momentlerini kullanarak, yineleme yapmadan model parametrelerini kestirmektir [8].

Lemma 1: İkinci ve üçüncü derece momentler model parametrelerine aşağıdaki gibi bağlıdır:

$$\mathbb{E}[x_3 \otimes x_1] = O A A \text{diag}(\pi) O^T \quad (5)$$

$$\sum_i \mathbb{E}[x_3 \otimes x_1 \otimes x_2 = i] \eta_i = \mathbb{E}[x_3 \otimes x_1 \otimes x_2](\eta) \quad (6)$$

$$= O A \text{diag}(O^T \eta) A \text{diag}(\pi) O^T$$

Burada η standart taban vektör kümesi \mathcal{E}^D 'den seçilmiştir.

Lemma 2: Öyle iki $U \in \mathbb{R}^{D \times K}$ ve $V \in \mathbb{R}^{D \times K}$ matrisi alalım ki $(U^T O A)$ ve $(V^T O \text{diag}(\pi) A^T \text{diag}(A \pi)^{-1})$ matrisleri tersinir olsun. $D \geq K$ koşulu sağlanırsa,

$$B(\eta) := (U^T \mathbb{E}[x_3 \otimes x_1 \otimes x_2](\eta) V) (U^T \mathbb{E}[x_3 \otimes x_1] V)^{-1} \quad (7)$$

$$= (U^T O A) \text{diag}(O^T \eta) (U^T O A)^{-1}$$

Burada U ve V için bir tercih $\mathbb{E}[x_3 \otimes x_1]$ matrisinin ilk K (sırasıyla) sağ ve sol tekil vektörlerdir. Buna göre, denklem (7)'deki $B(\eta) \in \mathbb{R}^{K \times K}$ matrisinin özdeğerlerinin $O^T \eta$ vektörünün elemanları, özvektörlerinin ise $U^T O A$ matrisinin sütunları olacağına dikkat çekelim. $B(\eta)$ tamamen gözlemler ışığında hesaplayabileceğimiz bir matristir, nitekim $\mathbb{E}[x_3 \otimes x_1 \otimes x_2](\eta)$ ve $\mathbb{E}[x_3 \otimes x_1]$ veriden hesaplanamın mümkün olduğu ampirik momentlerdir. Bu gözlemler doğrultusunda SMM parametreleri (\hat{O}, \hat{A}) öğrenilmesi için Algoritma 1 kullanılabilir.

Algorithm 1 SpektralSMM

Girdi: Gözlem dizisi $x_{1:T}$

Çıktı: Kestirilmiş Parametreler: \hat{O}, \hat{A} .

1. Veri $x_{1:T}$ 'den $\mathbb{E}[x_3 \otimes x_1]$ ve $\mathbb{E}[x_3 \otimes x_1 \otimes x_2]$ matrislerini kestir.
 2. Tekil değer ayrışımı $\mathbb{E}[x_3 \otimes x_1] = U \Sigma V^T$ ile U ve V matrislerini öğren.
 3. Özdeğer-özvektör ayrışımı $B(e_1) = R \text{diag}(O^T e_1) R^{-1}$ ile \hat{O} matrisinin ilk satırını ve $R = \widehat{U^T O A}$ matrisini öğren.
 4. $d \in \{2, \dots, D\}$ için $\hat{O}_{d,:} = R^{-1} B(e_d) R$.
 5. $\hat{A} = (U \hat{O})^{-1} R$
-

$B(\eta)$ matrisinin özdeğerlerinin sırası özvektörlerinin sırasına bağlı olduğuna dikkat çekelim. 1. adım haricindeki adımlarda R 'nin kullanılmasının sebebi, özdeğer sırası tutarlılığını korumaktır. Son olarak, burada, ampirik momentleri kullanırken bütün zaman anlarındaki veri $x_{1:T}$ 'yi kullanmamız mümkündür. Buradaki varsayım, saklı durum değişkenleri h_t Markov zincirini bir kararlı durum dağılımına yakınsadığıdır [7], [8].

III. SMM KARIŞIMI

A. Olasılıksal SMM karışım modeli

Karışım modelleri, gözlemlenen veriyi birden fazla dağılımın konveks kombinasyonu olarak modellemeyi sağlar. Veri $y_{1:N}$ için standart bir karışım modeli aşağıdaki gibi yazılabilir:

$$z_i | c \sim \text{Discrete}(c) \quad (8)$$

$$y_i | z_i, \theta_{1:M} \sim \prod_{m=1}^M p(y_i | \theta_m)^{[z_i=m]} \quad (9)$$

Burada $z_i \in \{1, \dots, M\}$ veri noktaları y_i 'lerin ait oldukları topakları belirten gösterge değişkeni, θ_m ise m 'inci topağın parametreleridir. $c \in \Delta^{K-1}$ ise karışım oranlarını veren olasılık dağılımıdır. Örneğin, yaygın olarak kullanılan Gauss karışım modelinde $p(y_i | \theta_m) = \mathcal{N}(x_i; \mu_m, \sigma_m^2)$ olarak alınır. Bu

çalışmada $p(y_i|\theta_m) = \sum_{h_{1:T,i}} \prod_{t=1}^T p(x_{t,i}|h_{t,i})p(h_{t,i}|h_{t-1,i})$ saklı markov modeli olarak alınmıştır.

B. Beklenti-embüytme algoritması ile SMM karışımı öğrenme

Parametrik bir SMM karışımı beklenti-embüytme algoritmasıyla öğrenilebilir [4]. Örnek olarak, ayırık gözlem dağılımı olan bir SMM için beklenti-embüytme algoritması aşağıdaki gibidir.

- Beklenti adımı:

$$w_{im} = \frac{c_m p(y_i|\theta_m)}{\sum_{m=1}^M c_m p(y_i|\theta_m)} \quad (10)$$

- Embüytme adımı:

$$O_{lu}^m = \frac{\sum_{i=1}^N w_{im} \sum_{t=1}^{T_i} [x_t = l] \mathbb{E}_m^i([h_t = u])}{\sum_{i=1}^N w_{im} \sum_{t=1}^{T_i} [x_t = l'] \mathbb{E}_m^i([h_t = u])} \quad (11)$$

$$A_{uv}^m = \frac{\sum_{i=1}^N w_{im} \sum_{t=1}^{T_i} \mathbb{E}_m^i([h_t = u][h_{t-1} = v])}{\sum_{i=1}^N w_{im} \sum_{t=1}^{T_i} \mathbb{E}_m^i([h_t = u'][h_{t-1} = v])} \quad (12)$$

$$\pi_u^m = \frac{\sum_{i=1}^N w_{im} \mathbb{E}_m^i([h_1 = u])}{\sum_{i=1}^N w_{im}} \quad (13)$$

Beklenti adımında i 'nci dizi y_i 'nin m 'inci topakta olma olasılığı w_{im} olarak hesaplanır. Embüytme adımında, m 'inci topak parametreleri O^m , A^m , π^m bütün saklı durum dizileri $h_{1:T}$ 'nin m 'inci topak için yeterli istatistikleri $\mathbb{E}_m([h_t = u])$ ve $\mathbb{E}_m([h_t = u][h_{t-1} = v])$ 'nin w_{im} ile ağırlıklı toplamı alınarak güncellenir. Diğer bir deyişle her yinelemede yeterli istatistikleri N tane dizi ve M topak için hesaplamamız gerekir. Dolayısı ile ortalama uzunluğu T , saklı durum sayısı K olan diziler için E beklenti-embüytme iterasyonu yaklaşık $\mathcal{O}(K^2 T N M E)$ hesap yüküne sahiptir (Yeterli istatistikler kanı yayılımı (belief propagation) algoritması ile $K^2 T$ zamanda hesaplanabilir.) Burada sunulan algoritmanın yumuşak topaklandırma olduğuna dikkat çekelim. Kesin topaklandırma yapmak için beklenti adımını, $m^* = \operatorname{argmax}_m p(y_i|\theta_m)$, $w_{i,m^*} = 1$, $w_{i,\bar{m}^*} = 0$ olarak değiştirmek yeterli olacaktır. Bu durumda yeterli istatistikler sadece topaktaki diziler kullanılarak hesaplanabilir. Bu algoritmanın karmaşıklığı ise yaklaşık olarak $\mathcal{O}(K^2 T N E)$ mertebesindedir.

C. Spektral öğrenme kullanarak SMM topaklandırma

Bu bölümde, SMM topaklandırmak için, spektral öğrenmeye dayanan yeni bir yöntem öneriyoruz. Bölüm II-C'de anlatılan spektral öğrenme yöntemine göre, belirli koşullar altında, bir SMM'yi sadece gözlemlenebilir momentleri kullanılarak öğrenebiliriz. Önerdiğimiz algoritma, bölüm III-B'de önerilen kesin topaklandırma algoritmasındaki parametre kestirme (embüytme) adımını spektral öğrenme yöntemi kullanarak yapmaktadır. Böylelikle, her yinelemedeki parametre kestirme adımı N adet T uzunluklu dizi üstündeki kanı yayılımı algoritması yerine, M adet tekil değer ayrışımı, M adet özdeğer ayrışımı ve $M \times L$ adet matris çarpımına dönüşmektedir. Dikkat edileceği üzere bu hesap yükü dizi sayısından bağımsızdır çünkü dizilerin ampirik istatistiklerini, algoritmanın başında bir kere çıkarmak yeterlidir. Algoritma 2'de önerilen yöntem özetlenmiştir.

Algorithm 2 Spektral yöntem ile K-SMM

```

for  $i = 1 \rightarrow N$  do
     $\mathbb{E}[x_3^i \otimes x_1^i \otimes x_2^i]$ ,  $\mathbb{E}[x_3^i \otimes x_1^i]$  istatistiklerini topla;
end for

Topak üyelik değişkenleri  $z_{1:N}$ 'yi rastgele ilklendir
for  $e = 1 \rightarrow \text{maxiter}$  do
    for  $k = 1 \rightarrow K$  do
         $\theta_k \leftarrow \text{SpektralSMM}(\{y_n | \forall n, z_n = k\})$ 
    end for

    for  $i = 1 \rightarrow N$  do
         $z_i = \operatorname{argmax}_k p(y_n | \theta_k)$ 
    end for
end for

```

Algoritma 2 içinde SpektralSMM parametre kestirimi algoritmasının (Algoritma 1) 2. adımını istatistikleri başta bir kez toplayarak atlayabiliriz. Topakların gözlemlenebilir momentleri $\mathbb{E}[x_3 \otimes x_1 \otimes x_2]$ ve $\mathbb{E}[x_3 \otimes x_1]$, $z_{1:N}$ aidiyet değişkenleri kullanılarak, algoritmanın ilk adımında çıkardığımız istatistikler sayesinde hızlı olarak hesaplanabilir.

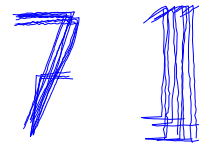
IV. DENEYSEL SONUÇLAR

A. Oyuncak Problem

Oyuncak problem olarak basit şekil gezinmelerini topaklandırmaya çalıştık. Deneklerden bilgisayar faresi ile çeşitli şekiller, örneğin, L, 7, L, ve V, çizmelerini istedik ve bu şekillerin uzamsal koordinatlarını kaydettik. Her bir sınıftan 10'ar gezeğe üretilmiştir. İki boyutlu x-y koordinat gezinmelerinin zamana göre birinci türevleri, gözlem modeli iki boyutlu yönbağımsız Gauss dağılımı olan SMMler ile modellenmiştir. Önerdiğimiz algoritma ile alınmış örnek sonuçlar Şekil 1 ve 2'de verilmiştir. Görüleceği üzere algoritma hata yapmaksızın gezinmeleri ayırt etmektedir.



Şekil 1. L ve V harf gezinmelerinin topaklandırılması.

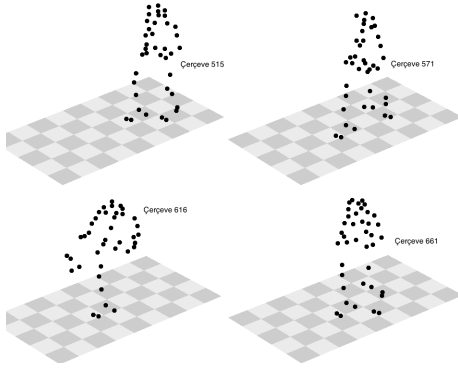


Şekil 2. 7 ve 1 rakam gezinmelerinin topaklandırılması.

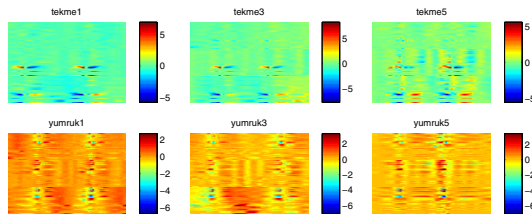
B. Devininim Yakalama Verisi

Önerdiğimiz yöntemi gerçek bir veri üstünde de denedik. Bu amaçla, HDM05 insan devinimleri (MoCap: Motion Capture) veritabanını [12] kullandık. Bu veri kümesinde insan vücudunun 40 farklı nirengi noktasının zamana bağlı

3B koordinatları verilmiştir. Şekil 3'de tekme atma sınıfına örnek bir dizi gösterilmektedir. Deneylerimizde, tekmeleme ve yumruklamadan oluşan iki sınıf edim dizisini topaklandırdık. Kısaca, her denek ileri doğru iki tekme ya da yumruk atmakta ve toplamda, her sınıf için ortalama çerçeve sayısı 500 olan 10'ar video dizisi bulunmaktadır. Bu iki edim birbirinden farklı gibi düşünülse de, bazı video dizilerinde sınıflar arası benzerlik yüksek olması, örneğin, ileri doğru yumruk atarken deneğin ayaklarını da hareket ettirmesi, ve farklı kamera açıları bu problemi zorlu kılmaktadır. Algoritmaya, 3B koordinatların zamana göre birinci türevini girdi olarak verdik. Durum sayısı $K = 3$ olan SMM kullandık. Gözlem modeli olarak yön bağımsız çok boyutlu Gauss dağılımı aldık. Tablo I'de "tekmeleme" ve "yumruklama" sınıflarından oluşan verikümesi üstünde elde edilen başarımlar verilmiş, bir döngü için geçen süreler karşılaştırılmıştır. Şekil 4'te ise aynı veri üstünde topaklandırılmış üç örnek dizi verilmiştir. Burada performans değerlendirmesi için yöntemleri rastgele ilklendirerek 10 defa başlattık ve ortalama başarımları (Tablo I, 2. sütun), ulaştıkları en yüksek başarımları (Tablo I, 3. sütun), bir döngü için geçen zamanı (Tablo I, 4. sütun) ve yakınsama için gerekli ortalama yineleme sayısını (Tablo I, 5. sütun) kaydettik. Beklenti-enbüyütme algoritmasının, parametre kestiriminde başarılı olması için parametrelerin iyi ilklendirilmesi gerektiğinden, denemelerin çoğunda en yüksek başarıma ulaşamamaktadır. Dolayısı ile ortalama başarımları daha düşüktür. Spektral yöntem ise, iyi ilkleme gerektirmedikinden bu problemle karşılaşmamakta ve en yüksek performans daha fazla ulaşmaktadır. Parametre ilklendirmesi gerektirmemesinin yanı sıra, spektral yöntemin esas üstünlüğü yineleme süresinin beklenti en-büyütme algoritmasına göre düşük olmasıdır. Ortalamada yakınsama için gereken adım sayısı daha fazla olmasına rağmen ortalama toplam sürede EM1'den yaklaşık 2 kat, EM2'den ise yaklaşık 3 kat hızlıdır.



Şekil 3. HDM05 [12] veritabanından örnek "tekmeleme" dizisi.



Şekil 4. "Tekmeleme" ve "yumruk" dizilerinin topaklandırılması sonucu elde edilen üç örnek dizi. Burada yatay eksen zaman, düşey eksen ise nirengi noktası indeksidir.

Tablo I. SPEKTRAL YÖNTEM VE BEKLENTI-ENBÜYÜTME ALGORİTMALARININ TEKMELEME VE YUMRUK SINIFLARINDAKİ 10'AR DİZİ ÜSTÜNDEN KARŞILAŞTIRILMASI. EM1: BEKLENTI-ENBÜYÜTME ALGORİTMASI KESİN TOPAKLANDIRMA, EM2: BEKLENTI-ENBÜYÜTME ALGORİTMASI YUMUŞAK TOPAKLANDIRMA.

	Ortalama Başarım(%)	Max Başarım(%)	Döngü süresi* (s)	Ortalama yakınsama
EM1	70	100	7.5	3.2
EM2	73	100	12	2.9
Spektral	76	100	3.1	3.8

*Gerçekleştirim: 3.33 GHz dual core CPU, 4 GB RAM, Yazılım: MATLAB

V. VARGILAR

Bu çalışmada SMM topaklandırmak için yeni bir algoritma önerdik. Bu algoritma, beklenti-enbüyütmeyle dayalı topaklandırma yaklaşımdan hesap yükü bakımından daha hafif olmakla beraber, başarı açısından karşılaştırılabilir sonuçlar vermektedir. İlerideki çalışmalarda, spektral öğrenme yöntemlerini diğer SMM topaklandırma yaklaşımlarına uyarlamaya çalışacağız. Örneğin, [6]'deki spektral topaklandırmaya dayalı yaklaşımda, parametre kestirme adımı, spektral öğrenme ile gerçekleştirilebilir. Ayrıca, yeni topaklar açmamızı sağlayan, parametrik olmayan Bayesçi yaklaşımlarda karşılaşılan hesaplaması güç integraller spektral öğrenme yardımı ile hesaplanabilir.

KAYNAKLAR

- [1] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition (1989)," *Proceedings of the IEEE*, pp. 257–286, 1989.
- [2] Y. Qi, J. Paisley, and L. Carin, "Music analysis using hidden markov mixture models," *Signal Processing, IEEE Transactions on*, vol. 55, no. 11, pp. 5209–5224, nov. 2007.
- [3] P. Smyth, "Clustering sequences with hidden markov models," in *Advances in neural information processing systems (NIPS)*, 1997.
- [4] A. Jonathan, S. Sclaroff, G. Kollios, and V. Pavlovic, "Discovering clusters in motion time-series data," in *Computer Vision and Pattern Recognition (CVPR)*, 2003.
- [5] T. Oates, L. Firoiu, and P. R. Cohen, "Clustering time series with hidden markov models and dynamic time warping," in *In Proceedings of the IJCAI-99 Workshop on Neural, Symbolic and Reinforcement Learning Methods for Sequence Learning*, 1999, pp. 17–21.
- [6] T. Jebara, Y. Song, and K. Thadani, "Spectral clustering and embedding with hidden markov models," in *European Conference of Machine Learning (ECML)*, 2007.
- [7] D. Hsu, S. M. Kakade, and T. Zhang, "A spectral algorithm for learning hidden markov models a spectral algorithm for learning hidden markov models," *Journal of Computer and System Sciences*, no. 1460-1480, 2009.
- [8] A. Anandkumar, D. Hsu, and S. Kakade, "A method of moments for mixture models and hidden markov models," in *Conference of Learning Theory*, 2012.
- [9] A. Anandkumar, R. Ge, D. Hsu, S. Kakade, and M. Telgarsky, "Tensor decompositions for learning latent variable models," 2012.
- [10] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in neural information processing systems (NIPS)*. MIT Press, 2001, pp. 849–856.
- [11] A. Parikh, L. Song, and E. Xing, "A spectral algorithm for latent tree graphical models," in *International Conference of Machine Learning (ICML)*, 2011.
- [12] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, and A. Weber, "Documentation mocap database hdm05," Universität Bonn, Tech. Rep. CG-2007-2, June 2007.