

YÜZ TANIMA YÖNTEMLERİNİN SIKIŞTIRILMIŞ ARŞİVLERDE BAŞARIMI

Mustafa Ersel Kamaşak ve Bülent Sankur
Sinyal ve İmge İşleme Laboratuvarı (BUSIM)
Boğaziçi Üniversitesi Elektrik Elektronik Mühendisliği
{kamasak, sankur}@busim.ee.boun.edu.tr

Özetçe

Sıkıştırma yöntemleri yeniden oluşturulan nesnelere hatayı en aza indirmek için tasarlanmıştır. Bundan dolayı, bu yöntemlerin sınıflandırma için gerekli bilgileri ne şekilde etkilediğini incelemeye çalıştık. Bu sıkıştırma algoritmalarının yüz tanıma yöntemlerinden “Özyüzler” ve “Fisher yüzleri” üzerindeki etkileri ve yüzlerin, tanıma başarımını çok etkilemeden, ne kadar sıkıştırılabileceği incelenmiştir.

1 GİRİŞ

Yüz tanıma araştırmaları giderek çoğalan pratik uygulama alanlarından dolayı ilgi çekmektedir. Uygulama alanlarına birkaç örnek, makina-insan arayüzü, giriş-çıkış kontrolü, ATM makinalarında güvenlik ve imge veritabanlarında arama olarak sıralanabilir.

İçerikleri giderek büyüyen yüz arşivlerini, depolanma sığası ve iletimdeki bant genişliği gibi kısıtlamalardan dolayı sıkıştırılmış olarak saklamak bir zorunluluk haline gelmiştir. Dolayısıyla, sıkıştırılmış ortamda yüz tanıma giderek önem kazanmaktadır. Böylece, gerek varolan tanıma algoritmalarının sıkıştırılmış verilerle başarımı, gerekse sıkıştırılmış nesnelere üzerinde yeni işleme ve tanıma yöntemlerinin geliştirilmesi güncel bir araştırma konusu olmuştur.

Bu bildiriye, sıkıştırma yöntemlerinin, doğrusal yüz tanıma yöntemlerine olan etkisi ve tanıma başarımını çok düşürmeden sıkıştırılabilme limitleri incelenmiştir.

2 YÜZ TANIMA YÖNTEMLERİ

Günümüzde birçok yüz tanıma yöntemleri geliştirilmiş olup [2], bunlar şu şekilde sınıflandırılabilir:

- *İlinti (correlation)* en basit yüz tanıma yöntemlerindedir. Yeni bir yüz ile veri tabanındaki yüzleri ilintiliyerek kimlik belirlemeye çalışır. Yüz imgelerindeki aydınlanma koşullarına, ölçeğe ve çevresel etkenlere karşı çok duyarlı olan bu yöntemin aynı zamanda işlem yükü de çok fazladır [1].
- *Doğrusal yöntemler* doğrusal ayrıştırma ile yüzsel özniteliklerin sınıflandırılmasına dayanır. “Özyüzler” (eigenfaces) [6] ve “Fisher yüzleri” [5] bu yöntemi kullanır. Hızlı oldukları ve kolay eğitilebildikleri için tercih edilirler.
- *Saklı Markov Serileri* ses işlemede uzun süredir kullanılan tek boyutlu “Saklı Markov Serilerinin” [10] yüz işleme için uyarlanması ile elde edilmiştir.
- *Uyarlanabilir dalgacık süzgeçleri* yüz imgelerinin uyarlanabilir dalgacıklar ile ayrıştırılmasına dayanır. Bu dalgacıklar yüzsel özniteliklerin elde edilmesi için gerekli, problem ve veriye özgü, tabanların elde edilmesini sağlar [7].

Bu çalışmada yüz tanıma algoritmaları olarak “özyüzler” ve “Fisher yüzleri” ele alınacaktır.

2.1 Özyüzler

Özyüzler yöntemi, yüz imgelerinin Karhunen-Loeve dönüşümü ile az boyutlu uzaya doğrusal izdüşümlerinin saçılımının en yükseğe çıkarılmasına dayanır. Bu yöntemde, yüzlerin n boyutlu uzaydan, m boyutlu öznitelik uzayına ($m \ll n$) izdüşümleri hesaplanır. x_k k'ncü yüz imgesinin tarama sırasına göre bir vektör olarak dizilimini ifade etsin. y_k ise aynı imgenin m özvektöründen oluşan, diğer bir deyişle m tane özyüzün ağırlıklandırılmış toplamından oluşan vektör olsun.

“ y_k ” vektörleri aşağıdaki doğrusal dönüşüm ile bulunur.

$$y_k = W^T x_k \quad x_k \in \mathcal{R}^n \quad \text{ve} \quad y_k \in \mathcal{R}^m \quad (1)$$

Bu ifadede $W \in \mathcal{R}^{n \times m}$ birimdik dikeçlerden (orthonormal columns) oluşan bir matristir. Eğer kovaryans matrisi S_T , eğitim kümesindeki yüzlerin sayısı N ve ortalama yüz imgesi μ olmak üzere,

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (2)$$

şeklinde tanımlanırsa, en az katsayıyla enerjinin en büyük kısmını ifade edecek dönüşüm matrisi W_{pca} şu şekilde bulunur:

$$W_{pca} = \arg \max_W |W^T S_T W| \quad (3)$$

Ana bileşenler analizini (Principal component analysis) gerçekleştiren W_{pca} matrisinin kolonları, saçılım matrisinin en yüksek özdeğerli m özvektörüdür. Yeni bir yüzü tanımak için, bu yüzün, W uzayındaki izdüşümü bulunur ve veri tabanında izdüşümleri daha önceden hesaplanmış yüzlerle karşılaştırılır. Yeni yüz, bu uzaydaki izdüşümü en yakın kişi olarak tanınır.

$$k = \arg \min_k \|Y_{test} - Y_k\|$$

Bu yöntem, nispeten büyük yüz arşivlerinden FERET yüz arşivi üzerinde test edilmiştir [4]. Değişen aydınlatma şartlarında %96, değişen bakış açılı yüzlerde %85, değişen boyutlardaki yüzlerde %64 başarımlıdır. Özyüz yönteminde kovaryansa yolaçan tüm etmenler, yani hem sınıflar-arası saçılımlar, hem de sınıf-içi saçılımlar da hesaba katılmış olur. Oysa sınıf-içi farklar, sınıflandırma işlemi için faydasız hatta zararlıdır.

2.2 Fisher yüzleri

Bu yöntemde ise Karhunen-Loeve dönüşü yerine “Fisher doğrusal ayrıştırma” dönüşümü kullanılır. Bu dönüşümde saçılma yol açan etmenlerin temsili yerine sınıflar arası saçılımın, sınıf-içi saçılımına oranı, yani sınıflandırılabilirlik en büyütülür [5].

Sınıflar-arası saçılımı, S_B ve sınıflar-içi saçılımı, S_W ile tanımlarsak ve c toplam sınıf sayısı olmak üzere sınıf ortalamasını ve nüfusunu, sırasıyla, μ_i ve N_i ile gösterirsek:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad S_W = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

Dönüşüm matrisi şu şekilde bulunur

$$W_{opt} = \arg \min_W \frac{|W^T S_B W|}{|W^T S_W W|} \quad (4)$$

$$= [w_1 \ w_2 \ \dots \ w_m]$$

Burada w_i , S_B ve S_W matrislerinin en yüksek değere sahip m özvektöründen biridir. Aynı zamanda en fazla $c - 1$ sıfırdan farklı özdeğer olacağından, m 'in üst değerinin $c - 1$ olacağı açıktır.

$$S_B w_i = \lambda_i S_W w_i \quad i = 1, 2, \dots, m$$

S_W 'nin kertesesi (rank) $N - c$ 'den az olduğu için W_{opt} doğrudan bulunamaz. Dolayısıyla imge ilk önce “özyüzler” yöntemi ile $N - c$ boyuta indirgenir. Daha sonra indirgenmiş uzayda “Fisher doğrusal ayrıştırma” yöntemi kullanılır.

$$W_{opt} = W_{pca} W_{fld} \quad (5)$$

burada

$$W_{pca} = \arg \max_W |W^T S_T W|$$

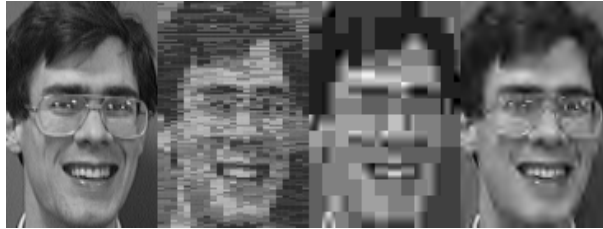
$$W_{fld} = \arg \max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$

Bu yöntemin Yale yüz arşivi [11] üzerindeki başarımı %99.4'dir. Bu arşiv, değişik aydınlatma koşullarında çeşitli yüz ifadeleri ve yüz takıları ile toplanmıştır. Aynı arşiv üzerinde “Özyüzler” metodunun başarımı sadece %80'dir [5].

3 SIKIŞTIRMA YÖNTEMLERİ

Kullanılan birçok sıkıştırma algoritmasından, aşağıdaki yöntemlerin en sık kullanılan ve diğer yöntemleri temsil eden bir küme oluşturduğuna inanıyoruz.

- *Vektör nicemleme* alt seviyeli imge işleme yöntemleriyle ortaklaşa kullanılabilceği ve kodlama işlemi, sınıflandırma başarımı ile bağdaştırılabileceğinden dolayı incelenmiştir [8] .
- *JPEG* birçok rakibi olmasına rağmen neredeyse sıkıştırma standardı haline geldiğinden dolayı incelenmiştir [8].
- *Dalgacıklar* JPEG'in en yakın takipçisi olduğundan ve kodlama için sıkça kullanılmaya başlandığından dolayı incelenmiştir. Burada ”embedded zero tree coding” algoritmasına dayanan SPIHT yöntemi kullanılmıştır [8, 9].



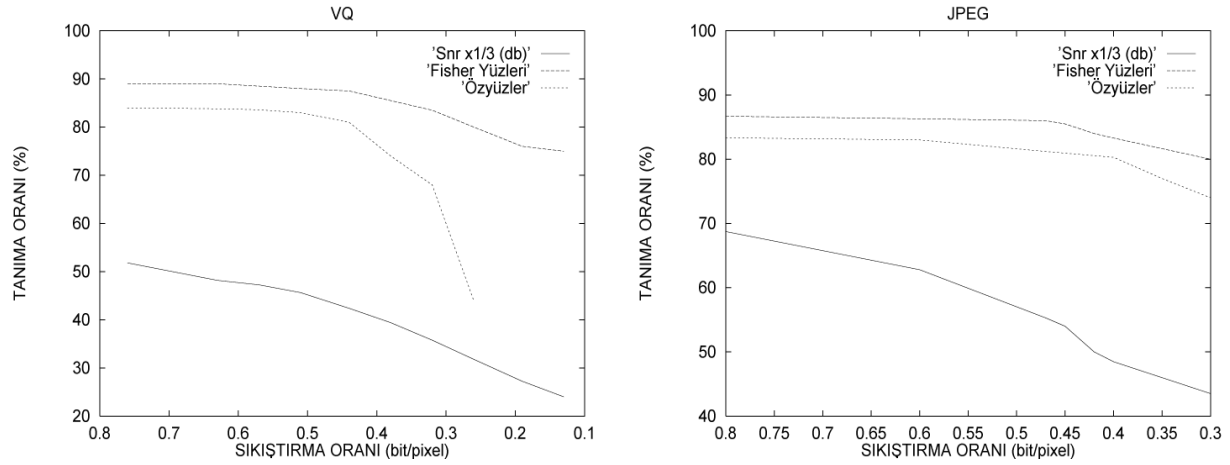
Şekil 1: Soldan sağa sırayla, orijinal ve 0.4 bit/piksel oranında vektör nicemleme, JPEG, ve SPIHT kullanılarak kodlanmış yüzler

Bu yöntemlerle yapılan düşük hızlı kodlama, imgelerde çeşitli bozukluklara sebep olabilir. Bu bozukluklardan birkaçı bloklanma, bulanma, ayrıntı kaybı, gölgelenme, ayrıntıların basamaklanmasıdır. Örneğin VQ ve JPEG, imgelerde bloklanma ve ayrıntı basamaklanmasına sebep olur, buna karşın alt-band kodlama da bulanma ve halkalanma bozukluklarına sebep olur.

4 DENEY SONUÇLARI

Deneylelerimizde 40 kişinin 10 farklı açıdan çekilmiş ve 92x112 çözünürlükte taranmış imgelerinden oluşan ORL yüz veri tabanını [3] kullanıldı. Tüm deneylerde, insan yüzleri, özyüzler ve Fisheryüzleri yöntemlerini kullanarak 15 boyutlu ”yüz uzayına” indirildi ve tanıma işlemi bu uzayda gerçekleştirildi.

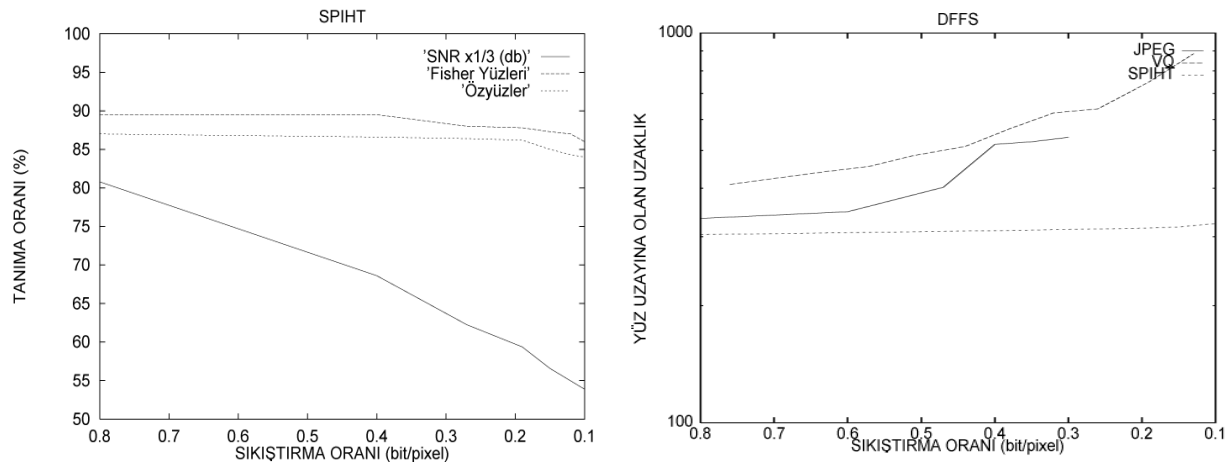
Sıkıştırılmamış yüz imgelerinde, özyüzler yönteminin başarımı %88 iken, Fisherface yönteminin başarımı %89.5'dir. Bu iki skorun birbirine yakın çıkmasının nedeni, bu veri tabanındaki imgelerde aydınlanma koşulları ve artalan gürültüsü fazla değişmemesi, ve yüzlerin boyut ve yerlerinin neredeyse aynı olmasıdır.



Şekil 2: (a) Vektör nicemleme (b)JPEG yöntemiyle sıkıştırılmış yüzlerin tanıma yüzdesi ve işaret-gürültü oranı

Şekil 2a, 2b ve 3a da sırasıyla, VQ, JPEG ve SPIHT algoritmaları ile 10 ila 40 kat sıkıştırılmış arşivlerdeki tanıma başarımları verilmiştir.

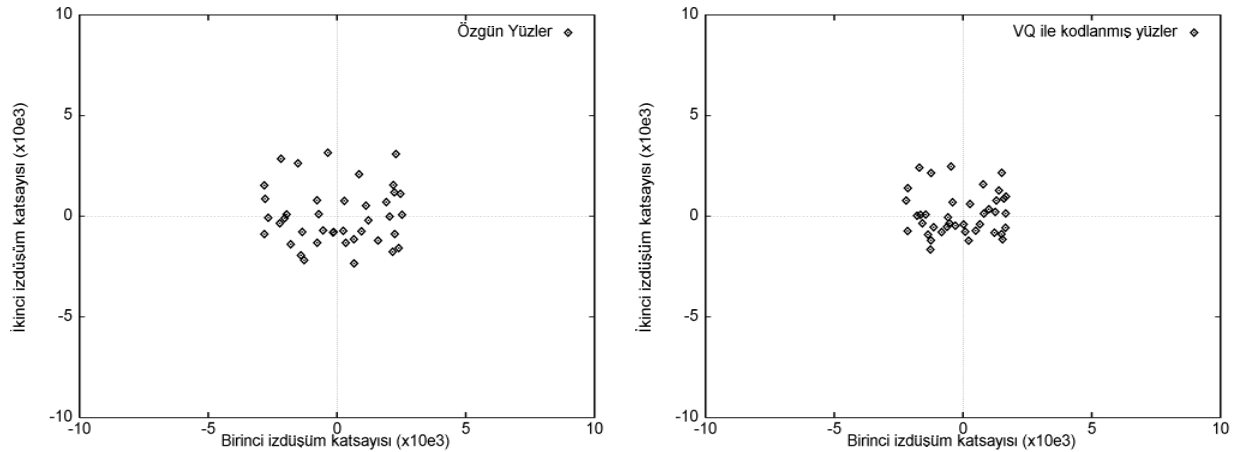
Her üç algoritma da 20 katı kadar, yani 0.4 bit/piksel oranına kadar sıkıştırılmaya dayanmaktadır. Bu hızın ötesinde VQ ve JPEG, bir kırılma noktasına varmakta, diğer bir deyişle çok hızlı bir performans düşümüne sahne olmaktadır. Oysa SPIHT 40 katı kadar sıkıştırmaya, yani 0.2-0.1 bit/pixel hızlarına kadar dayanmaktadır. Bu arada Fisher yüzleri her kodlama hızında ve yönteminde özyüzlere göre % 2-5 kadar daha iyi bir başarımla sonuçlanmaktadır.



Şekil 3: (a) Dalgacık yöntemiyle sıkıştırılmış yüzlerin tanıma yüzdesi (b) Sıkıştırılmış yüzlerin, yüz uzayına olan uzaklıkları

Altband kodlama ise kullanılan kodlama yöntemleri arasında en başarılısı olup 0.2 bit/piksel hıza kadar kodlanan yüz imgelerini performans yitirmeden tanıyabilmiştir.

Kodlama yöntemlerinin, yüz özneliklerini ne kadar koruyabildiğini görmek için “yüz uzayından uzaklık” [6] ölçüsünü ve özdeğer saçılımını kullandık. Şekil 3b’den de görüldüğü gibi, altband kodlama yöntemleri artan sıkıştırmayla yüz uzayından daha az uzaklaşmakta, oysa JPEG ve VQ giderek daha az yüze benzemektedir.



Şekil 4: Sırasıyla sıkıştırılmamış ve vektör nicemleme ile 0.4 bit/piksel hızında kodlanmış yüzlerin ilk iki özyüz üzerindeki izdüşümlerinin saçılımı

Şekil 4’den görülebileceği üzere, sıkıştırılmış imgelerin, oluşturulan yüz uzayındaki izdüşümlerinin saçılımları daralmıştır. Bu da yüzlerin taşıdığı ayırt edici bilginin azaldığını göstermektedir.

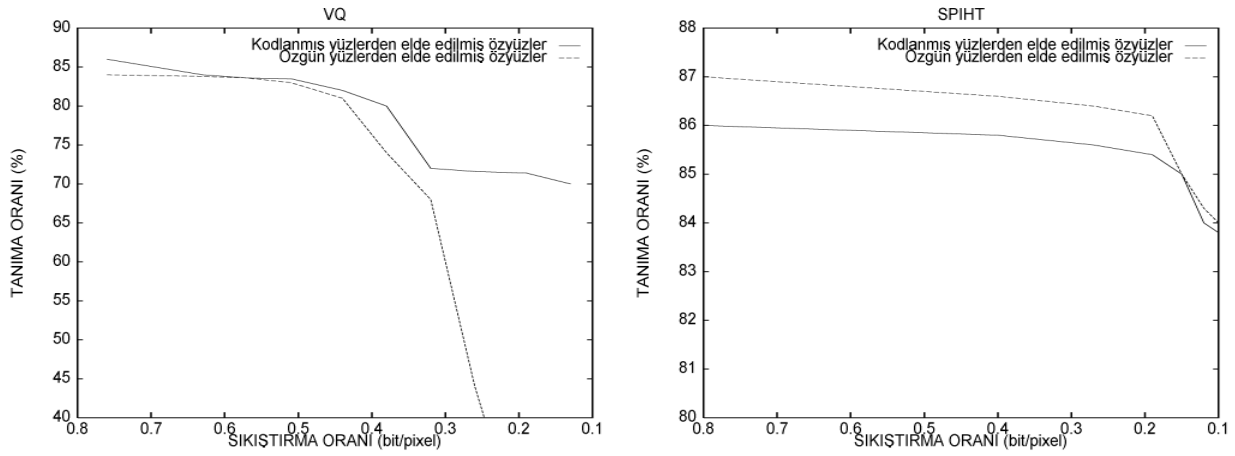
Son olarak, özyüzlerin özgün yüzler yerine sıkıştırılmış yüz imgelerinden oluşturmanın tanıma başarımını ne şekilde etkilediğini inceledik. JPEG ve vektör nicemlemede 0.4 bit/piksel hıza kadar sıkıştırılmış imgerlerden elde edilen özyüzlerden % 1-2 daha iyi sonuç vermekte, oysa dalgacıklarda, sıkıştırılmamış imgelerden elde edilen özyüzlerden daha iyi sonuçlar vermektedir (Şekil 5)

5 SONUÇLAR

Yüz imgeleri 0.4 bit/piksel hızına kadar, tanıma başarımları çok azalmadan vektör nicemleme, JPEG veya dalgacıklar yöntemlerinden biriyle kodlanabilir. Eğer imgelere 0.4 bit/piksel hızın altında kodlama yapılacaksa, bu hızlarda dalgacık yöntemleri, vektör nicemleme ve JPEG’den daha başarılı olmakta ve 0.2 bit/piksel hızına kadar kodlanabilmektedir. Ayrıca dalgacık yöntemi yüzün özneliklerini diğer yöntemlerden daha iyi koruduğu anlaşılmıştır. Tüm kodlama yöntemleri için Fisheryüzleri tanımda özyüzlerden % 3-5 daha başarılı olmuştur.

Kaynakça

- [1] Brunelli R., Poggio T., “*Face Recognition: Features versus Templates*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, s. 1042-1052, Vol 15, Ekim 1993.
- [2] Kirby M., and Sirovich L. ”*Application of Karhunen - Loeve Procedure for the Characterization of Human Faces.*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 12, s. 103-108, 1990.
- [3] ORL Yüz Arşivi, <http://www.cam-orl.co.uk/facedatabase.html>
- [4] Pentland A., Starner T., Etcoff N., Masoiu A., Oliyide O. and Turk M., “*Experiments with Eigenfaces*”, Looking at People Workshop IJCAI’93, Chamberry, France, Ağustos 1993.



Şekil 5: Sıkıştırılmış ve özgün imgelerden elde edilen özyüzlerin karşılaştırılması (a) VQ (b) SPIHT

- [5] Peter N., Hespanha J., Kriegman D., “*Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 19, No 7, Temmuz 1997.
- [6] Turk M., Pentland A., “*Eigenfaces for Recognition*”, Journal of Cognitive Neuroscience, Vol 3, s. 71-86, No 1, 1991.
- [7] Phillips J., “*Matching Pursuit Filters Applied to Face Identification*”, IEEE Transactions on Image Processing, Vol 7, No 8, Ağustos 1998, s. 1150-1164.
- [8] Rabbani M. J., *Digital Image Compression Techniques*, SPIE Optical Engineering Press, 1991.
- [9] Said A., Pearlman W. A., “*A New and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees*”, IEEE Transactions on Circuits and Systems for Video Technology, Vol 6, s. 243-250, Haziran 1996.
- [10] Samaria F., Fallside F., “*Face segmentation for Identification and Feature Extraction using Hidden Markov Models*”, Image Processing: Theory and Applications, 1993.
- [11] Yale Yüz Arşivi, <http://cvc.yale.edu>