

parlaklık bilgisi kolaylıkla çıkarılabilir. RGB bileşenleri aşağıdaki formül ile YCbCr bileşenlerine çevrilebilir.

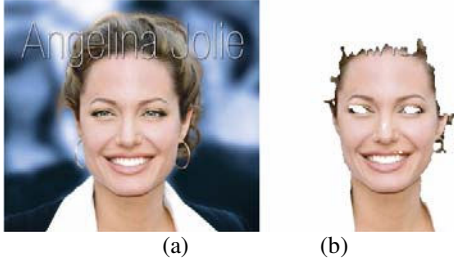
$$\begin{aligned} Y &= 0,299R + 0,587G + 0,114B \\ Cb &= -0,169R - 0,332G + 0,500B \\ Cr &= 0,500R - 0,419G - 0,081B \end{aligned} \quad (1)$$

Ten bulma işleminde Cb ve Cr değerleri kullanılmaktadır. Cb ve Cr bileşenlerinin maksimum ve minimum değerleri hesaplanarak bu değerler arasında kalan pikseller ten olarak işaretlenir. Cb ve Cr bileşenlerinin minimum ve maksimum değerlerine, bu bileşenlerin ortalama değerine (mean) ve standart sapmalarına göre karar verilir [6].

$$\begin{aligned} Cb_{min} &= Mean_Cb - Std_Cb * f \\ Cb_{max} &= Mean_Cb + Std_Cb * f \\ Cr_{min} &= Mean_Cr - Std_Cr * f \\ Cr_{max} &= Mean_Cr + Std_Cr * f \end{aligned} \quad (2)$$

Bu çalışmada, Mean_Cb = -11.1051, Std_Cb = 4.3568, Mean_Cr = 22.9265, Std_Cr = 3.9479 ve f = 3 olarak kullanılmıştır. Sabit f değeri değiştirilerek sonuçlar gözlenebilir.

Şekil 2'de örnek resme ilişkin bulunan ten bölgesi sonucu gösterilmiştir.



Şekil 2. Orijinal resim (a), Ten bulunmuş resim (b)

Ten bulma işleminde ten olarak işaretlenen, ancak yüz olamayacak kadar küçük olan alanların elenmesi için etiketleme yöntemi kullanılmış ve böylece tarama alanı daraltılmıştır. Bulunan muhtemel ten bölgelerinin yüze ait olması, dolayısıyla yüz tespiti Yapay Sinir Ağları yardımıyla gerçekleştirilmektedir.

2.2. Yapay Sinir Ağları

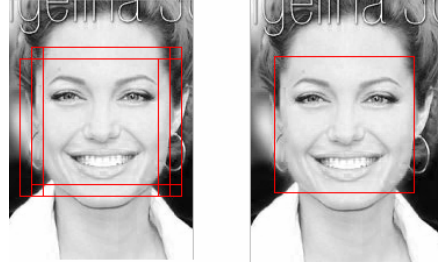
Yapay sinir ağları (YSA) birbirine bağlı çok sayıda işlemci biriminin oluşturduğu bir sistemdir. İnsan beyni deneyimlerinden faydalanarak öğrenir [7]. YSA, biyolojik sinir ağları model alınarak geliştirilmiştir. Biyolojik sinir hücresinin becerisine sahip olmak için yapay sinir hücrelerinde uygun dört temel fonksiyon modellenmiştir. Tanıma sistemlerinde genellikle çok katmanlı ağlar (ÇKA) kullanılmaktadırlar. Bunun nedeni, ÇKA'nın doğrusal olmayan problemleri başarıyla çözmesidir. Fakat çözebileceği problemlerin sınırları çok iyi bilinmemektedir.

İlk olarak yapay sinir ağı oluşturulur ve eğitime başlanır. Yapay sinir ağına, eğitimde kullanılmak üzere, çıkış değeri "1" olacak şekilde belirlenen yüz resimleri ve çıkış değeri "0" olacak şekilde belirlenecek olan yüz olmayan resimler yüklenmektedir. Yapay sinir ağındaki hücrelerin aktivasyon fonksiyonlarının belirlenmesi kullanıcıya bağlıdır. Bu çalışmada, aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmıştır.

$$y_j = \frac{1}{1 + e^{-x_j}} \quad (3)$$

Burada x_j , ara katmanın j. ara sinirinin çıkış değerini göstermektedir.

Yapay sinir ağını eğitmek için kullanılan örnekler çok önemlidir. Ağ, ne kadar iyi eğitilir ise o kadar iyi sonuç alınır. Çalışmada, yapay sinir ağına eğitimi 25x25 boyutlarda resimler ile yaklaşık 300 civarında yüz ve 600 civarında yüz olmayan toplam 900 örnek ile gerçekleştirilmiştir, ara katman sayısı 10, momentum 0.2 ve öğrenme katsayısı ise 0.1 alınmıştır. YSA yardımıyla ten bölgeleri bulunmuş resim üzerinde yüz bulma işlemi gerçekleştirilir. Böylece sadece ten bölgelerinde arama gerçekleştirilerek zamandan kazanılır ve yanlış sonuçlar verebilecek bölgeler elemine edilmiş olur. Bunun sonucunda, bulunan yüz bölgeleri çerçevelenerek gösterilir ve birkaç çerçevenin üst üste bindiği görülür. Aynı yüze ilişkin seçilmiş bir çok bilgidan ortak tek bilginin elde edilmesi için Kutucuk-Birleştirme (Box-Merge) algoritması kullanılmıştır[6].



Şekil 3. Tek çerçeveye sınırlandırılan yüz görüntüsü

Şekil 4'te, ten işleme ile güçlendirilmiş YSA yardımıyla çeşitli açılarda ve sayılarda yüz görüntüsü olan resimlerden yüz tespiti sonuçları gösterilmiştir.



Şekil 4. Bulunan yüzler

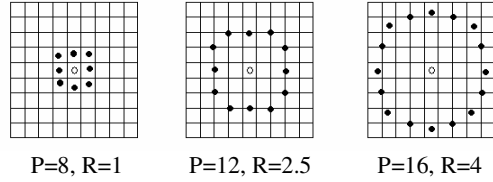
3. Yüz İfadesinin Değerlendirilmesi

Çalışmada 2. aşama olan resimden yüz ifadesinin değerlendirilmesi de iki kısımdan oluşmaktadır. Öncelikle bulunan yüze ait özellik vektörü çıkarılır ve daha sonra bu özellik vektörü sınıflanarak yüzün hangi ifadeye ait olduğu sonucuna ulaşılır. Özellik çıkarma işlemi için Yerel İkili Örnekler (Local Binary Patterns - LBP) yaklaşımı kullanılmıştır.

3.1 Yerel İkili Örnekler (Local Binary Pattern - LBP)

Yerel İkili Örnekler (LBP) operatörü, ilk olarak Ojala vb. tarafından ortaya çıkarılmıştır [8]. LBP özellikleri, çeşitli uygulamalarda örneğin doku sınıflamasında ve bölütlenmesinde, yüz tanımda, yüz ifadesi tanımda vs. çok iyi performans göstermiştir. LBP operatörü, gri seviye ölçüsü olarak tanımlanır.

Görüntüdeki her piksel için yerel komşulukları merkez piksele göre eşiklendirilerek bir ikili kod üretilir. Bu gri seviye operatörünün, değişmeyen örneklerin sabit bir rotasyon kümesi içerdiği görülmüştür [9]. Örneklerin dairesel gösteriminde sınırlı sayıda geçişler vardır. Komşuluklarına göre farklı LBP operatörleri tanımlanabilir. Genel olarak $LBP_{P,R}$ üç farklı dairesel simetrik komşuluklar kümesi ile tanımlanabilir, P komşu sayısını, R ise örnekleme yarıçapını gösterir. Şekil 5'te çeşitli LBP operatörleri gösterilmiştir.



Şekil 5. Çeşitli dairesel LBPP,R operatörleri.

Çalışmada her pikselin 3x3 komşuları için $LBP_{8,1}$ yapıları operator kullanılmıştır ve merkez piksele göre eşiklendirilerek LBP kodu hesaplanmıştır.

$$s(g_i - g_c) = \begin{cases} 0, & \text{eğer } g_i < g_c \\ 1, & \text{aksi takdirde} \end{cases} \quad (4)$$

Burada g_c , merkez pikselin gri seviye değerini göstermektedir. Bir özel dönüşüm fonksiyonu olan $s(g_i - g_c)$ ise (5) bağıntısına göre verilmektedir.

$$s(g_i - g_c) = \begin{cases} 0, & \text{eğer } g_i < g_c \\ 1, & \text{aksi takdirde} \end{cases} \quad (5)$$

$LBP_{8,1}$ operatörü için bakıldığında $2^8 = 256$ tane farklı LBP ikili kodu olmaktadır. Bütün LBP örneklerinin doku özelliğini iyi tanımlamadığı bilinmektedir. LBP, merkez pikselin komşularına karşı düzenliliğini aynen alır, fakat komşularının düzenliliğini aynen almaz [10]. Bu yüzden, örneklerdeki düzenliliği tanımlamak için bir düzenlilik ölçüsü U tanımlanmıştır. İkili dizi dairesel olduğunda 0-1 ya da 1-0 en fazla iki bit geçişi varsa LBP örneği düzenlidir denir. Doku görüntülerinde, (8; 1) komşulukta tüm örneklerin %90'nını, (16; 2) komşulukta ise %70'ni düzenli örneklerin oluşturduğu gözlenmiştir. 8-bitlik LBP kodu için 58 tane düzenli LBP kod örnekleri vardır, buna göre $256-58=198$ tane de düzensiz örnekler bulunmaktadır [11]. U_c , merkez piksel c'nin düzenlilik ölçüsü olarak ifade edilir ve aşağıdaki şekilde tanımlanır.

$$U_c = \sum_{i=0}^{P-1} |s(g_i - g_c) - s(g_{i-1} - g_c)| \quad (6)$$

$$LBP_{c,j} = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c), & \text{eğer } U_c \leq j, j=2 \\ P+1, & \text{aksi takdirde} \end{cases} \quad (7)$$

U_2 , (8; 1) rotasyonu komşulukta düzenli örnekleri gösterir. Bu çalışmada 3x3 ($LBP_{8,1}$) operatörü kullanılmıştır. Yöntemde 58 düzenli örneğin dışında kalan örnekler tek bir örnek altında etiketlenir ve böylece 256 yerine $58+1=59$ tane özellik vektörü elde edilir.

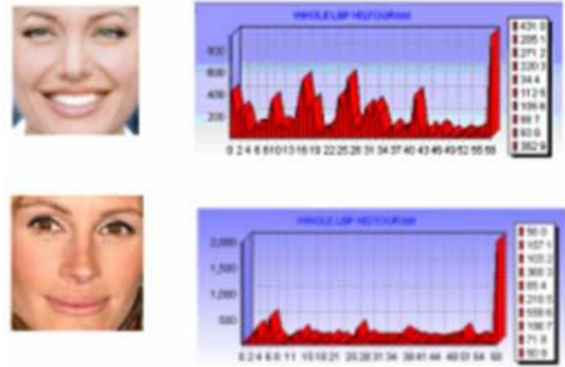
Etiketlenmiş görüntünün görüntünün $f(x,y)$, histogramı aşağıdaki şekilde tanımlanır.

$$H_i = \sum I(f(x,y) = i), \quad i=0, \dots, n-1 \quad (8)$$

n, LBP operator tarafından üretilen farklı etiketlerin sayısı ve I(A), bir özel dönüşüm fonksiyonu iken

$$I(A) = \begin{cases} 1, & \text{eğer } A \text{ doğru ise} \\ 0, & \text{eğer } A \text{ yanlış ise} \end{cases} \quad (9)$$

Bu histogram, kenarlar, noktalar ve köşeler gibi yerel bir çok küçük örneklerin tüm görüntüde dağılımı hakkında bilgi içerir. Bu çalışmada görüntülerin H_i histogramlarını incelendiğinde genel olarak, yüz ait resimlerin $H[58]$ 'de, yüz olmayan resimlerin ise $H[0], H[1], \dots, H[4]$ 'de en yüksek değerlere sahip oldukları gözlemlenmiştir.



Şekil 6. Çeşitli yüz ifadesi olan resimlere ait H_1 histogramları

Yüz resimleri, LBP histogramları ile etkili bir biçimde tanımlanabilen çok küçük örneklerin birleşimi olarak görülebilir. Tanıma işlemi için yerel bölgeler üzerindeki LBP histogramları daha güvenilir bir tanımlama sağlamaktadır. Bu nedenle, LBP özelliklerinin elde edilmesi için yüz resimleri eşit olarak küçük bölgelere bölünür. Bununla birlikte yüz resim boyunun ve pozisyonunun ($[+15^\circ, -15^\circ]$) sabit olması LBP için bir dezavantajdır, fakat iyi sonuçlar elde edilmesinde oldukça etkilidir [12]. Çalışmada, 100x100 boyutlarında resimler kullanılmış ve bu resimler 10x11 boyutlarında alt parçalara bölünmüştür. Bunun sonucunda 90 tane alt bölge elde edilmiş ve her bölge için 59 özellik tanımlandığından, tüm resim için 59X90 tane özellik vektörü elde edilmiştir. Bu vektöre LBP yerel histogramları denilir ve aşağıdaki gibi tanımlanır,

$$H_{i,j} = \sum_{x,y} I(f(x,y) = i), \quad (10)$$

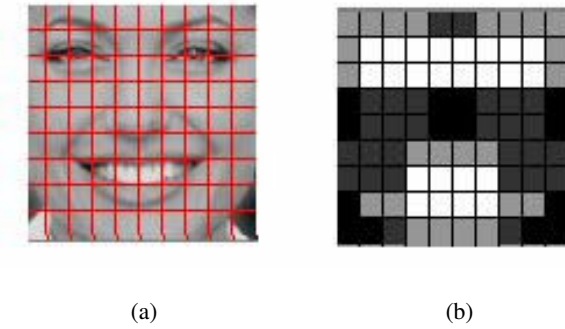
Burada $I(f(x,y) \in R_j), i=0, \dots, n-1; j=0, \dots, m-1$ olmaktadır. m bölge sayısını, n ise LBP operatörü tarafından üretilen 59 etiket sayısını ve R_j bölgeleri göstermektedir. Bu çalışmada, $m=90$ 'dır. LBP operatörünün bir avantajı da histogramların çok hızlı bir şekilde hesaplanabilmesidir.

3.2. Örnek Tanıma (Sınıflama)

Çalışmada psikolojik durum değerlendirilmesinde, kolaylığı ve başarısı açısından model eşleme (template matching) metodu seçilmiştir. Bunun için, öncelikle her ifadeye ait özellik vektörünün belirlenmesi gerekmektedir. Bu çalışmada, her ifadeye ait özellik vektörü, o ifadenin eğitim kümesindeki resimlerin özellik vektörlerinin ortalaması olarak belirlenmiştir. Tanıma işlemi için en yakın komşu sınıflaması kullanılmıştır. Giriş resmi, LBP histogramı en yakın model ile eşleştirilir ve sonuç üretilir.

3.2.1. Chi Karesi İstatistiği (Chi Square Statistics)

Giriş resminin, her ifadeye ait LBP histogramlarına uzaklığını ölçmek için model eşleme tabanlı Chi karesi istatistiği (X^2) yöntemi kullanılmıştır. Yüze ait ifadelerin, çoğunlukla göz ve ağız bölgeleri gibi belli bölgelerde bulunduğu ve bu bölgelerin yüz ifadesi tanımada daha yararlı bilgiler içerdiğini gözlemlenmiştir [13]. Bu yüzden, her yüz bölgesi için içerdiği bilginin önemine göre bir ağırlık belirlenebilir. Çalışmada belirlenen şablonda siyah kareler ağırlık olarak 4.0, koyu gri kareler 2.0, açık gri kareler 1.0 ve beyaz kareler 0.0 olarak alınmıştır.



Şekil 7. (a) Bölünmüş resim, (b) Ağırlıklandırılmış resim.

Minimum uzaklık o yüze ait ifadeyi vereceğinden, ağırlıklar önemli bölgelerden önemsiz bölgelere doğru artan sırada seçilmektedir. Ağırlıklandırılmış (X^2) istatistik ölçütü (11) bağıntısıyla belirlenmektedir.

$$X_w^2(S,M) = \sum_{i,j} w_j \cdot \frac{(S_{i,j} - M_{i,j}^c)^2}{S_{i,j} + M_{i,j}^c} \quad (11)$$

Burada S örneğe, M ise modele ait histogramları ve w_j , bölgelere ait ağırlıkları göstermektedir. Her ifade (model) için ağırlıklandırılmış Chi karesi hesaplaması yapılarak minimum uzaklık bulunur ve giriş örneğinin hangi ifadeye ait olduğu belirlenir.

3.2.2. Minimum Uzaklık (Minimum Distance)

Minimum uzaklık sınıflayıcısı, bilinmeyen görüntü verilerini, çoklu özellik uzayında görüntü ve sınıf mesafelerini küçültürerek sınıflamak için kullanılmaktadır. Bu mesafe bir benzerlik indeksi olarak tanımlanır. Böylece minimum uzaklık maksimum benzerlikle hemen hemen aynıdır.

$$D^2 = \sum_{ij} (S_{ij} - M_{ij})^2 \quad (12)$$

3.2.3. Değiştirilmiş Minimum Distance (Changed Minimum Distance)

Burada, minimum uzaklıktaki farkların kareleri yerine mutlak değerleri alınmıştır. Giriş ve model örnekleri arasındaki fark ne kadar az ise o kadar çok yakındırlar mantığı ile farkların mutlak değerleri alınıp toplandığında ve daha sonra da bu toplama kare kökü alma işlemi uygulandığında iyi sonuçlar elde edilmiştir.

$$D^2 = \sum_{i,j} |S_{ij} - M_{ij}| \quad (13)$$

Sınıflama metodlarının birleştirilmesinde, her metod sonucu 0-1 değerleri arasına normalize edilip, metodlar ağırlıklandırılarak sonuçlar toplanmıştır.

$$\sum_{i=0}^n K_i * P_i \quad (14)$$

Burada, n metod sayısını, K_i normalize edilmiş metod sonucunu ve P_i ise metodların yüzde ağırlıklarını göstermektedir.

4. Tartışma

Yüz ifadesi tanıma ile ilgili anlatılan metodlar karşılaştırıldığında tablodaki sonuçlar elde edilmiştir.

Tablo 1. Farklı metodların karşılaştırılması

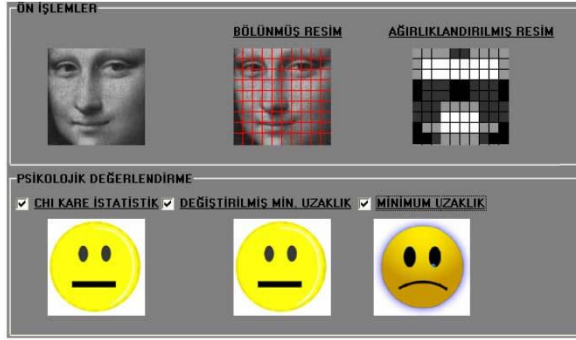
Kaynak	Yöntem	Kısıtlamalar	Doğruluk
Shinohora [1]	HLAC + Fisher ağırlık haritası	Yüzün el ile kesilmesi	% 69.4
Lyons[2]	Wavelet + PCA+LDA	El ile seçilmiş 34 sabit nokta	% 75
Önerilen yöntem	LBP+ağırlıklandırılmış Chi karesi ve uzaklık Sınıflandırıcıları	Yüzü modele uygun kesmek	% 77

Tablo 1'e bakıldığında LBP operatörünün özellik çıkarma işleminde oldukça etkili olduğu görülmektedir. Fakat her yöntemin getirdiği bir kısıtlama vardır. Bu da yüz ifadesi tanıma işleminin zorluğundan kaynaklanmaktadır. Eğer test örneğinin ifadesi, eğitim setindeki örneklerden oldukça farklı ise yanlış sonuçlar çıkabilir, bu nedenle sağlam ve iyi bir eğitim seti oluşturulursa sistemin doğruluğu arttırılır.

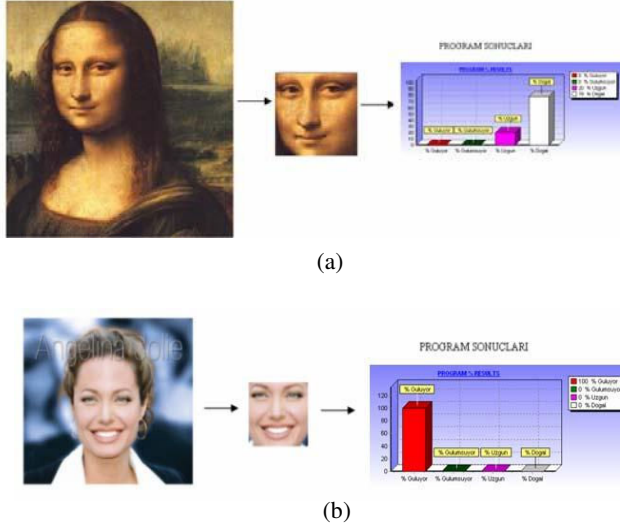
5. Sonuçlar

Çalışmada, çeşitli yüz ifadelerinin tanınması üzerinde durulmaktadır. . *Gülen, Doğal ve Üzgün* olmak üzere 3 farklı yüz ifadesi incelenmiş, sınıflandırılarak tanınmaya çalışılmıştır. Bu amaçla deri analizi ve Yapay Sinir Ağları kullanılarak giriş bilgisinden önce yüz tespiti gerçekleştirilmiştir. Sonra elde edilen yüz örneği üzerinde özellikler çıkarılarak sınıflandırma sonucunda tanıma işlemi gerçekleştirilmiştir. Tespit edilen yüzler üzerinde özellik çıkarma işlemi için kullanılan LBP

yöntemi, Gabor dalgaları ile karşılaştırıldığında özellik çıkarma bakımında daha hızlı olduğu ve daha az boyutlu uzayda yer aldığı görülür. Yüz ifadelerini sınıflamak için model eşleme tabanlı Chi kare istatistik, minimum uzaklık ve değiştirilmiş minimum uzaklık metodları kullanılmıştır. Belirgin ifadeler için üç sınıflama metodu da aynı sonucu verirken, birden fazla ifadenin gizli olduğu resimlerde farklı sonuçlar verebilmektedir. Bu nedenle, çıkan sonuçların modele göre yüzdeleri ifadelendirilerek, sınıflama sonuçları birleştirilmiştir. Bu şekilde daha sağlam sonuçlar elde edilmiş olur. Bununla birlikte, bir veritabanı programı hazırlanarak insanlardan psikolojik durum değerlendirilmesi yapılması istenmiştir ve sonuçlar veritabanında tutulmuştur, daha sonra bu sonuçlar program sonuçları ile karşılaştırılarak değerlendirilmiştir.



	Chi-kare istatistik	Minimum Uzaklık	Değiştirilmiş min. uzaklık
Doğal	12141	247.072	95.681
Gülen	12445	252.687	97.226
Üzgün	12199	242.082	96.891



Şekil 8. Çeşitli yüz ifadelerine ilişkin program sonuçları
(a) %78 normal - %20 üzgün, (b) % 100 gülüyor

Gerçekleştirilen sistem yüz ifadelerine ilişkin olarak kesin kararlar verdiği gibi (örneğin %100 gülme vb.) dereceli sonuçlar da verebilmektedir, Şekil 8 a,b. Elde edilen bu deneysel sonuçlar sistemin insana benzer davranışlarının modellenmesi gibi yorumlanabilir.

6. Kaynakça

- [1] Shinohara, Y. and Otsu, N., "Facial Expression Recognition Using Fisher Weight Maps", *IEEE Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 499-504.
- [2] Lyons, M., Budynek, J., and Akamatsu, S., "Automatic Classification of Single Facial Images", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, 1999, pp. 1357-1362.
- [3] Zhang, Z., "Feature-based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-layer Perceptron", *Pattern Recognition and Artificial Intelligence* Vol. 13, 1999, pp. 893-911.
- [4] Zhang, Z., Lyons, M., Schuster, M., and Akamatsu, S., "Comparison Between Geometry-based and Gabor-Wavelet based Facial Expression Recognition Using Multi-layer Perceptron", *Proc. 3rd Int. Conf. Automatic Face and Gesture Recognition*, 1998, pp. 454-459.
- [5] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory: Eurocolt '95*, pages 23-37. Springer-Verlag, 1995.
- [6] Inseong Kim, Joon Hyung Shim, and Jinkyu Yang, "Face Detection", *Stanford University*.
- [7] Ho S.Y., Huang H.L., "An analytic solution for the pose determination of human faces from monocular image", *Pattern Recognition Letters, Elsevier*, 1998
- [8] Ojala, T., Pietikinen, M., and Harwood, D., "A comparative study of texture measures with classification based on featured distribution", *Pattern Recognition*, vol. 29, no. 1, 1996.
- [9] Stan Z. Li, Chunshui Zhao, Meng Ao, Zhen Lei, Center for Biometrics and Security Research & National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences
- [10] Ojala, T., Pietikinen, M., and Menp, T., "Multi-resolution grayscale and rotation invariant texture classification with local binary patterns", *IEEE PAMI*, vol. 24, no. 7, July 2002.
- [11] Tsolmongerel Orkhonselenge, "Texture based segmentation of remotely sensed imagery for identification for geological units", February, 2004.
- [12] Caifeng Shan, Shaogang Gong and Peter W. McOwan, "Conditional Mutual Information Based Boosting for Facial Expression Recognition", *Queen Mary, University of London, London, E1 4NS, UK*
- [13] Caifeng Shan, Shaogang Gong and Peter W. McOwan, "Robust Facial Expression Recognition Using Local Binary Patterns", *Queen Mary, University of London, London, E1 4NS, UK*