

Göz Bebeği Merkezi İçin Artırımlı Kestirim Yöntemi

Iterative Estimation of The Eye Pupil Center

Necmeddin Said Karakoç, Şamil Karahan ve Yusuf Sinan Akgül
GTÜ Bilgisayarla Görme Lab, <http://vision.gyte.edu.tr/>, Bilgisayar Mühendisliği Bölümü
Gebze Teknik Üniversitesi, Kocaeli, 41400, Türkiye
{said.karakoc,samilkarahan}@gmail.com, akgul@gtu.edu.tr

Özetçe—Göz bebeği merkezinin kestirimini yapan ticari sistemlere alternatif olarak geliştirilen tak çalıştır kamera sistemleri üzerinde son zamanlarda bilgisayarla görme alanında sıklıkla çalışılmaktadır. Bu çalışmada, herhangi bir özel donanıma ihtiyaç duymadan Destek Vektör Regresörleri (DVR) kullanılarak göz bebeği merkezi tespiti yapılmıştır. DVR eğitimi için göz bölgesinden elde edilen gradyan yönü tabanlı öznitelik vektörü ile bu bilginin etiketi olarak göz bebeği merkezinin uzaklığı kullanılmıştır. Test aşamasında, kayan pencere yöntemiyle elde edilen görüntü parçasına göre regresörün ürettiği göz bebeği merkezine olan uzaklık bilgisinin kullanılmasıyla skor görüntüsü oluşturulmuştur. Oluşturulan bu skor görüntüsünün zirve noktaları bulunarak nihai göz bebeği merkezlerinin kestirimi yapılmıştır. Teste başlamadan önce göz tespit edici ile belirlenen alana göre kestirimi yapılan göz bebeği merkezlerinin iyileştirilmesi için artırımlı bir yaklaşım kullanılmıştır. Bu yaklaşıma göre, merkez nokta değişmeye kadar bir önceki aşamada elde edilen sonuç bilgisiyle belirlenen alanda test işlemleri tekrarlanmıştır. Önerilen sistem, BioID veri kümesi üzerinde yapılan test sonuçlarında literatürdeki en gelişmiş tekniklerin başarısının üzerinde bir performans göstermiştir.

Anahtar Kelimeler —Göz Bebeği Merkezi Tespiti; Destek Vektör Regresörleri; Yönlü Gradyan Histogramu.

Abstract— Off the shelf camera based eye pupil center detection has been very popular among computer vision community for the recent years. We propose an accurate and robust regressor based pupil center estimation method without any specialized hardware. The method trains a Support Vector Regressor using HOG features against the Euclidean distance between the center of the train patches and the ground-truth pupil center. On the test stage, we employ a sliding window approach to produce a score image that contains the regressor estimated distances to the pupil center. We select the best center position among the candidate centers by fitting a second degree polynomial to the maximal score image positions. The detected locations are improved by using an iterative method that repeats the center finding operation until there is no change. We evaluate our method on the challenging BioID data set. The results of the experiments are overall very promising and the system exceeds performance of the similar state of the art methods.

Keywords — Pupil Center Estimation; Support Vector Regression; HOG.

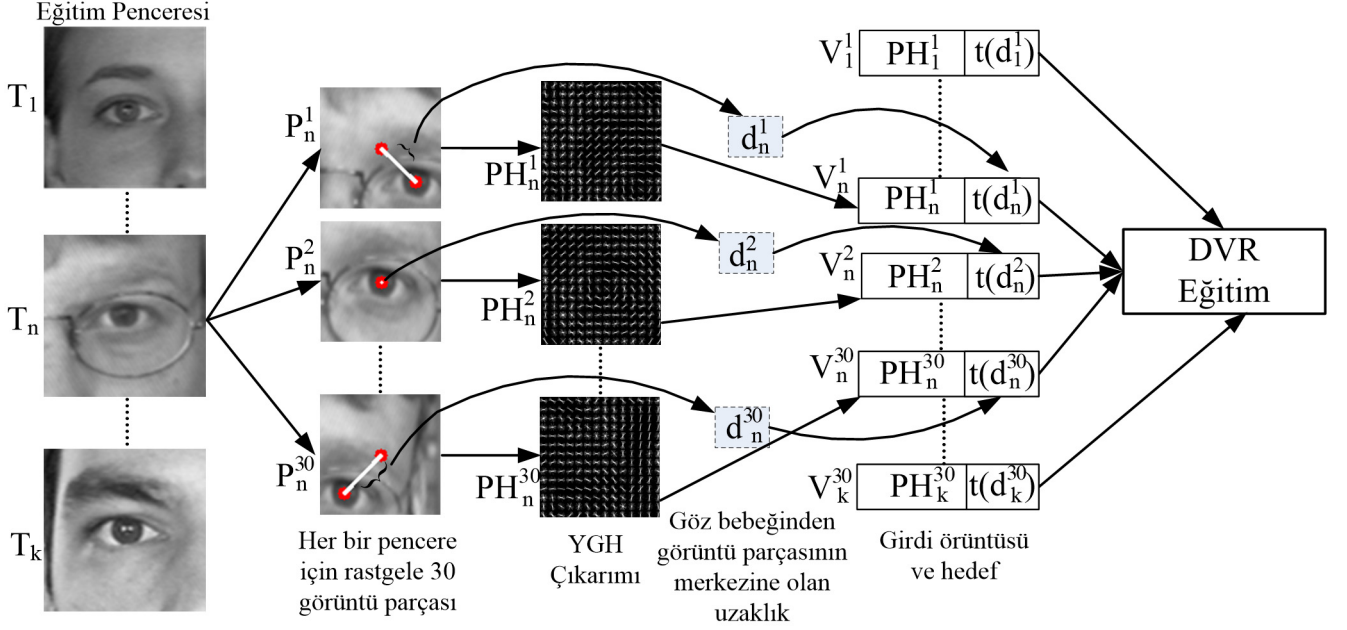
I. GİRİŞ

Göz bebeği merkezinin doğru bir şekilde tespit edilmesi göz takibi, insan-makina ara yüzleri, kullanıcının dikkat takibi ve engelli insanlar için göz ile kontrol edilebilir sistemler gibi birçok uygulamada kritik öneme sahiptir. Bu merkez noktanın tespit edilmesinde kullanılan yöntemlerden birisi de kafaya takılan özelleşmiş cihazlar veya birden fazla kızılötesi kamera kullanımınıdır. Bu yöntem göz bebeği merkezini doğru bir şekilde tespit etmesine rağmen; beraberinde pahalı, rahat kullanılamaması ve kalibrasyona gereksinim duyması gibi dezavantajlar taşımaktadır.

Son yıllarda sadece tek web kamerasına ihtiyaç duyan görünüm tabanlı göz bebeği merkezinin yerinin saptanmasına yönelik çalışmalar literatürde yer almaya başlamıştır. Bu yöntemler genel hatlarıyla model tabanlı, öznitelik tabanlı ve hibrit olmak üzere 3 ayrı kategoriye ayrılmaktadır [1]. Bunlardan ilki olan model tabanlı yöntemler, genellikle bir kümenin öznitelikliğini çıkararak sınıflandırma ya da regresyon yapar veya öğrenilmiş bir modeli göze uydurur. Markus vd. [2] merkezin kestirimi için rasgele regresör ağaç topluluğuna dayalı bir yöntem önermişlerdir. Bu yöntem MS Kinect [3] ile insan vücudunun parçalarını sınıflandırmada kullanılan yöntemle benzer bir şekilde göz bebeği merkezinden elde edilen rastgele noktalar üzerinde çalışmaktadır.

Bir diğer yöntemde ise öznitelik tabanlı metotlar makine öğrenmesi kullanmadan bilinen göz özelliklerini kullanarak (köşeler, kenarlar ve gradyanlar gibi) göz bebeğini merkezinin kestirimi yapmaktadır. Timm vd. [4] görüntü gradyanlarının analizine dayalı bir yaklaşım önermektedir. Bu yöntem için tanımlanan amaç fonksiyonu bütün gradyan vektörlerinin göz bebeği merkezinde kesiştiğini kabul eder. Bu yöntem de bizim gibi gradyan bilgisini kullanmasına rağmen bizim bu bilgiden model öğrenmemiz göz bakışının ve kafa pozisyonun değişimine karşı daha gürbüz olmamızı sağlamıştır.

Hibrid metotlar, öznitelik tabanlı yöntem ile belirlediği aday göz bebeği merkezi noktalarından en uygun olanı öğrenilen model tabanlı yöntemle seçmektedir. Valenti vd. [1] eş aydınlıklar eğrisi öznitelikliğini kullanarak aday göz bebeği merkezlerini elde etmektedir. Daha sonra SIFT özniteliklerini kullanarak oluşturduğu model sayesinde nihai merkez noktanın yerini kestirmektedir.



Şekil 1. Önerilen yöntemin eğitim aşaması.

Bu çalışmada, monoküler kameralardan elde edilen görüntüler üzerinde yer alan göz bebeği merkezinin doğru ve gürbüz bir şekilde bulunması için yeni model tabanlı bir yöntem önerilmiştir. Yapılan çalışma sonucunda göz bebeği görüntüsündeki gradyan açısı bilgisinin özgün bir imza özelliği taşıdığı gözlemlenmiştir. Bu imza bilgisi Yönlü Gradyan Histogramı (YGH) [5] kullanılarak elde edilmiştir. Genelde ikili sınıflandırma ile nesne tespitinde kullanılan YGH özneliği bu çalışmada, göz bebeği görüntüsünü içeren görüntü parçasından elde edilen YGH vektörünün öznitelik ve bu görüntü parçasının merkezinin göz bebeği merkezine olan piksel uzaklığının etiket olarak kullanılmasıyla Destek Vektör Regresör (DVR) [6] modeli oluşturulmuştur. Bu yaklaşım, parçaların göz bebeği merkezine olan uzaklığını sürekli değerler olarak sunması sayesinde merkeze ne kadar yaklaşıldığı konusunda anlamlı bilgiler vermektedir. Tahmin edilen aday göz bebeği merkezlerinin zirve noktalarının başlangıçta merkez olarak kabul edilen noktaya en yakın olanı yeni merkez nokta olarak işaretlenmiştir.

Halen devam etmekte olan çalışmamızdan [7] elde edilen merkez noktaları kullanarak arama penceresi boyutunun ve konumunun güncellenmesiyle sonuçların iyileştirilmesi için artırılmış bir yaklaşım uygulanmıştır. Buna göre, başlangıçta belirlenen göz merkezleri, artırılmış olarak iyileştirilmiştir. Her artırım sonucunda kestirimi yapılan göz bebeği merkezinin bir önceki merkeze olan uzaklığının eşik değerinden daha fazla değişmemesi durumunda artırım sonlandırılmıştır. Önerilen yöntem gözün uç noktalara bakması ve kafanın dönük olması durumlarında gürbüz ve doğru bir tahmin yapmaktadır.

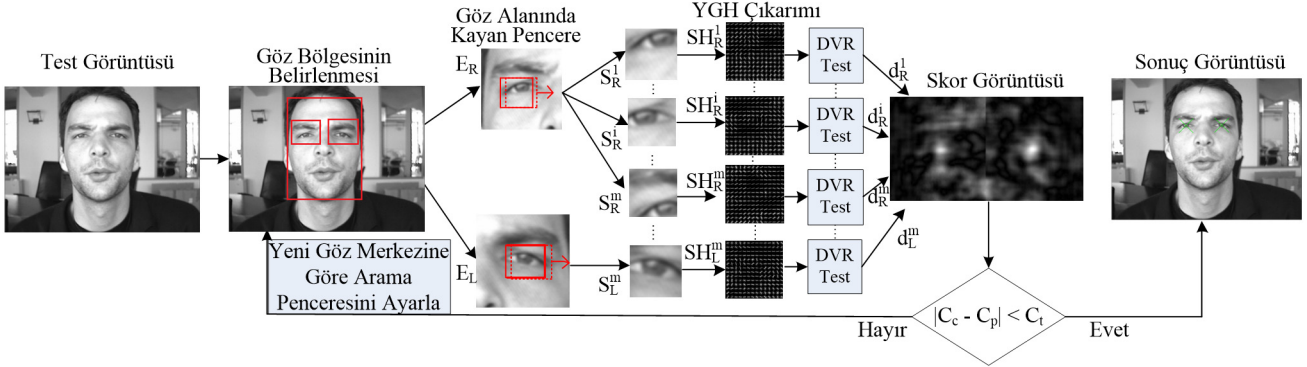
Bildirinin kalan kısmı şu şekilde oluşturulmuştur: 2. bölümde önerilen metod açıklanmakta; 3. bölümde deney

sonuçları açıklanmakta ve son zamanlarda bu konuda yayınlanan çalışmalarla karşılaştırmaları içermekte; 4. bölüm ise çalışmanın son değerlendirmesini yapmaktadır.

II. YÖNTEM

Önerilen method, eğitim ve test olmak üzere iki aşamadan oluşmaktadır. Eğitim için (Bknz. Şekil 1) göz merkezi etrafından rastgele k tane görüntü parçası (T_1, T_2, \dots, T_k) alınmıştır. Alınan bu parçaların boyutları gözün iris bölgesinin yaklaşık 3 katıdır. Göz merkezi etrafından elde edilen bu görüntü parçalarının her biri (T_n) eğitilirken çözünürlüğünün 96×96 olması sağlanmıştır.

Göz merkezi etrafından alınan görüntünün tam merkezinde göz bebeği içerdiğini varsayalım (Bknz. Şekil 1, P_n^n). f , görüntü parçalarından öznitelik çıkarımı yapan bir fonksiyon olsun. Amacımız, bu öznitelik bilgisini kullanarak P_r^n 'nin merkezinin göz bebeği merkezine olan Öklid uzaklığını veren bir $y(\cdot)$ fonksiyonu öğrenmektir. Bu fonksiyon P_r^n 'nin merkezinin göz bebeğinden uzaklaşmasıyla azalacak ve $y(f(P_n^n))$ ifadesinde ise en iyi değeri üretecektir. Bu çalışmada, $y(\cdot)$ fonksiyonun öğrenilmesi için DVR yöntemi kullanılmıştır. Bu modeli eğitmek için ihtiyaç duyulan girdi örüntüleri f fonksiyonu ile üretilen öznitelik vektörleridir ve bunun için YGH öznitelik çıkarımı kullanılmıştır. Bu regresör modelinin girdi örüntülerine karşı ihtiyaç duyduğu hedef bilgi için, P_r^n 'nin merkezinin gerçek göz bebeği merkezine olan Öklid uzaklığının Denklem 1'deki üstel fonksiyona girdi olarak girilmesi sonucu elde edilen değer kullanılmıştır. Bu fonksiyonun aldığı değer sıfıra yaklaştıkça, yani göz bebeği merkezi ile P_r^n merkezinin yakın olmasıyla daha yüksek değerler üretmektedir [8].



Şekil 2. Önerilen sistemin test aşaması.

$$t(d_r^n) = \begin{cases} e^{a(1 - \frac{d_r^n}{PS})} - 1 & \text{eğer } d_r^n < PS \text{ ise} \\ 0 & \text{diğer durumlarda,} \end{cases} \quad (1)$$

burada $a > 0$ ifadesi fonksiyonun göz merkezi yakınlarındaki büyüme hızını belirleyen bir sabittir. PS, eğitilecek görüntünün büyüklüğüdür. Eğer Öklid mesafesi bu değerden büyükse fonksiyon sıfır üretmektedir. Yapılan deneylerde t fonksiyonunun d_r^n ye göre daha iyi sonuçlar ürettiği gözlemlenmiştir. Şekil 2'de öğrenilmiş üstel fonksiyonun ürettiği değerlerin 0-255 arasına sıkıştırılmış değerlerinin oluşturduğu bir skor görüntüsü gösterilmiştir.

f fonksiyonu, eğitilecek olan P_r^n görüntüsünü alıp YGH vektörü PH_r^n 'yi üretmektedir. Son olarak V_r^n vektörü, YGH vektörü PH_r^n ile $t(d_r^n)$ hedefinin birleştirilmesiyle oluşturulmaktadır. Burada $n = 1, 2, \dots, k$ ve $r = 1, 2, \dots, 30$ değerlerini almaktadır. Önerilen bu yöntemde, sol gözler yatay hizada terse çevrilerek her iki göz için sadece bir regresör modelinin kullanılmasını sağlamıştır. Yapılan deneyler sonucunda, regresör için üçüncü dereceden polinom çekirdeğinin alternatiflerine göre daha iyi çalıştığı gözlemlenmiştir.

Test aşamasında (Bknz. Şekil 2), verilen bir I görüntüsünden öncelikle yüz ve göz bölgeleri yaygın kullanıma sahip Haar [9] yöntemiyle tespit edilmiştir. Dikdörtgen şeklinde tespit edilen bu alanın merkezi başlangıç göz merkezi olarak kabul edilmiştir. Bu merkez noktalara göre gözler arası mesafeden göz genişlik ve yükseklikleri, ve buna bağlı olarak regresyon pencere boyutu belirlenmiştir (Şekil 2, E_L and E_R). Bu pencere, 96x96 çözünürlüğüne ölçeklenerek kayan pencere yöntemiyle regresör sonuçlarına göre skor görüntüsü oluşturulmuştur (Bknz. Şekil 2). Skor görüntüsünde, merkez etrafındaki zirve noktaları tespit edilmiştir. Elde edilen bu zirve noktaları ikililerinden, aradaki uzaklıkları başlangıç merkez uzaklıklarına en yakın olanları seçilir. Yeni göz bebeği merkezi olarak işaretlenen noktaların (C_c) sağ ve sol göz için ayrı ayrı bir önceki merkezlere (C_p) olan uzaklığı belirlenen eşik değeri C_t ile kıyaslanmıştır. Bu kıyaslama sonucuna göre eğer değişim eşik değerinin

üzerinde ise, test aşaması bulunan yeni göz bebeği merkezlerini başlangıç olarak kullanılarak tekrarlanmıştır. Bu işlem, kıyaslama sonucunun eşik değerinden düşük oluncaya kadar devam edecektir. Bu şekilde artırimsal olarak sonuçların her aşamada iyileşmesi sağlanmıştır. Önerilen sistemin test aşaması Şekil 2'de görsel olarak anlatılmıştır.

III. DENEYLER

Yapılan çalışmanın deneyleri için göz bebeği merkezi tespiti yapan çalışmalarda sık olarak kullanılan ve çeşitli zorlu durumları içeren BioID veri tabanı [10] kullanılmıştır. Bu veri tabanı, farklı zamanlarda ve farklı yerlerde çekilmiş 24 farklı kişiye ait 1500'den fazla gri görüntüden oluşmaktadır. Veri tabanı gözlüklü kişiler, gözün kapalı olması, göz bölgesinde oluşan gölge efektleri, yansımadan dolayı görünmeyen göz bebeği vb. zor durumları içeren görüntüleri de içermektedir. 286x384 çözünürlüğe sahip olan bu görüntüler aynı kalitede ve düşük çözünürlüklü bir kamerayla çekilmiştir. Ayrıca, göz bebeği merkezlerinin görüntü üzerindeki konum bilgisi de veri kümesinde yer almaktadır. Deneyler sırasında toplamda 24 kişinin bulunduğu veri tabanındaki her bir 3 kişi, kalan 21 kişi ile eğitilerek test edilmiştir. Bütün kişilere ait görüntülerin test edilebilmesi için 8 farklı kombinasyonla testler tamamlanmıştır. Bu sayede test için belirlenen bir kişinin eğitim aşamasında herhangi bir görüntüsünün olmadığı garanti edilmiştir.

Elde edilmiş göz merkezlerinin doğruluğunu ölçmek için normalize edilmiş hata değeri hesaplanmaktadır. Jesorsky vd. [11] tarafından tanımlanan bu ölçüm şu şekilde hesaplanmaktadır:

$$e \leq \frac{\max(e_L, e_R)}{d} \quad (2)$$

burada e_L ve e_R sırasıyla sol ve sağ göz bebeği merkezinin gerçek yeri ile test aşamasında kestirilen yeri arasındaki Öklid mesafesini göstermektedir. d ise gerçekteki sol ve sağ göz bebeği merkezi arasındaki Öklid mesafedir. Tablo 1'de önerdiğimiz yöntemin normalize edilmiş sonuçlarının literatürdeki bilinen yöntemlerle kıyaslanması gösterilmiştir. Tablodan da görüleceği üzere, önerilen



Şekil 3. BioID veri kümesinden örnek test sonuçları. Son sütun sistemin başarısız olduğu görüntüleri göstermektedir. Artı işareti tam doğruyu gösterirken, çarpı işareti ise bizim kestirimini yaptığımız göz bebeğini göstermektedir.

çalışma bütün kategorilerde en iyi sonuçları elde etmiştir. Sistem test aşamasında, göz tespit algoritmasının her iki gözü de bulabildiği görüntüleri kullanmıştır. Diğer çalışmalar testlerini nasıl yaptıklarının detaylarını raporlamamıştır. Göz bölgesinin tespitinden oluşan hataların en aza indirilmesi için göz bebeğinin arandığı pencere diğer yöntemlere göre daha büyük seçilmiştir. Şekil 3’de önerilen yöntemin BioID veri kümesindeki bazı görüntüler üzerindeki test sonuçlar gösterilmiştir.

Yöntem	$e \leq 0.05$	$e \leq 0.1$	$e \leq 0.25$
Önerilen Çalışma	%97.5	%99.6	%99.9
Karakoc vd. [7]	%92.2	%97.7	%99.6
Markus vd. $p = 31$ [2]	%89.9	%97.1	%99.7
Tim vd. [4]	%82.5	%93.4	%98.0
Valenti vd. [1]	%86.1	%91.7	%97.9

Tablo 1. BioID veri kümesi üzerinde normalize edilmiş hataların diğer yöntemlerle karşılaştırılması.

IV. SONUÇLAR

Bu bildiride Destek Vektör Regresörleri kullanılarak görünüm tabanlı göz bebeği merkezi tespiti yöntemi önerilmiştir. Bu yöntem, aday göz görüntülerinin YGH öznelikleri çıkartılarak göz merkezi ile görüntü merkezi arasındaki uzaklığı regresör fonksiyonu sayesinde tahmin etmektedir. Kullanılan yöntem YGH öznelikleri çıkardığı için, göz görüntüleri iyi çözünürlükte olmasa bile gradyan açısından faydalanarak iyi sonuçlar elde edebilmektedir. Test aşamasında artırımı bir yöntemle her seferinde daha iyi göz bebeği merkezinin bulunması sağlanmıştır. Standart veri kümeleri üzerinde yapılan deney sonuçları, önerilen metodun başarısını göstermektedir. Bu çalışmanın devamında, yöntemin hızlanması için regresör sonuçlarına uygulanacak en dik tırmanma algoritmasıyla göz merkezine doğru daha hızlı yaklaşılması düşünülmektedir.

KAYNAKÇA

- [1] Valenti, R., Gevers, T., 2012. Accurate eye center location through invariant isocentric patterns. TPAMI 34, 1785-1798.
- [2] Markuš, N., Frljak, M., Pandžić, I. S., Ahlberg, J., & Forchheimer, R. (2014). Eye pupil localization with an ensemble of randomized trees. Pattern recognition, 47(2), 578-587.
- [3] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman and A. Blake, "Real-Time Human Pose Recognition in Parts from Single Depth Images," Communications of the ACM, vol. 56, no. 1, pp. 116-124, 2013.
- [4] Timm, F., Barth, E., 2011. Accurate eye centre localisation by means of gradients. In: Int. Conference on Computer Vision Theory and Applications. Vol. 1. pp. 125-130.
- [5] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.
- [6] H. Drucker, C. J. Burges, L. Kaufman, A. Smola ve V. Vapnik, "Machines, Support Vector Regression," Advances in neural information processing systems, 1997, pp. 155-161.
- [7] N. S. Karakoc, S. Karahan ve Y. S. Akgul, "Regressor Based Estimation Of The Eye Pupil Center," ICIIP, 2015, submitted.
- [8] Sironi, Amos, Vincent Lepetit, and Pascal Fua. "Multiscale Centerline Detection by Learning a Scale-Space Distance Transform." Conference on Computer Vision and Pattern Recognition (CVPR). No. EPFL-CONF-198153. 2014.
- [9] BioID Image Dataset (Accessed Jan 2015), <https://www.bioid.com/About/BioID-Face-Database>
- [10] V. Paul and M. J. Jones, "Robust real-time face detection," International journal of computer vision, vol 57, no. 2, pp. 137-154, 2004.
- [11] J. Oliver, K. J. Kirchberg and R. W. Frischholz, "Robust face detection using the hausdorff distance," in In Audio-and video-based biometric person authentication, 2001, pp. 90-95.