

Grafiksel Model Yardımıyla Göz Öznitelik Noktalarının Takibi

Eye Feature Point Tracking by Using Graphical Models

Serhan Coşar, Müjdat Çetin

Mühendislik ve Doğa Bilimleri Fakültesi
Sabancı Üniversitesi, Tuzla/İSTANBUL

serhancosar@su.sabanciuniv.edu, mcetin@sabanciuniv.edu

Özetçe

Bu makalede, göz öznitelik noktalarının takibi için istatistiksel bir yöntem önerilmektedir. Önerilen yöntemin amacı, gözlemlenen verilerin gürültülü ve/veya görüntüde yüzün önünde oluşabilecek engeller nedeniyle güvenilir olmadığı durumlarda bile öznitelik takibinin başarıyla yapılabilmesini sağlamaktır. Buradan yola çıkılarak kurulan grafiksel model, nokta hareketlerindeki zamansal bilgiye ek olarak noktalar arasındaki uzamsal ilişkilerle ilgili bilgileri de kullanmaktadır. Önerilen yöntem 2-B gri video dizileri üzerinde denenmiş, yaklaşımın mevcut tekniklere üstünlüğü gösterilmiştir.

Abstract

In this paper, a statistical method for eye feature point tracking is proposed. The aim is to track feature points even when the observed data are uncertain because of noise and/or occlusion. With this motivation, a graphical model that uses the spatial information as well as the temporal information between points is built. The proposed method is applied on 2D grayscale real video sequences as a real data application.

1. Giriş

Yüz öznitelik noktalarının takibi, örneğin göz, ağız, burun öznitelik noktaları, yüz ifade analizi, yüz ifadesi tanıma, insan-bilgisayar ara yüzü gibi konularda var olan basamaklardan biridir. Video dizileri üzerinde yapılan bu çalışmalar tipik olarak 3 bölümden oluşmaktadır: öznitelik çıkarımı, öznitelik takibi ve tanıma/algılama süreci. Öznitelik çıkarımı, imgelerdeki hareketleri tanımlayabilen, ayırt edici özellikleri olan noktaların bulunması işlemini içermektedir. Örneğin, yüz imgeleri için kaşların, gözlerin, ağzın, burnun bulunması ve/veya bu organların hareketlerindeki kritik noktaların bulunması. Bu aşamadan sonra süre gelen takip aşaması çıkarılan bu özniteliklerin dizi içindeki hareketlerinin takibinden oluşmaktadır. Son olarak takip edilen hareketler tanıma/algılama aşamasında değerlendirilerek yüz ifadesi olarak çıktıları alınmaktadır. Örneğin, mutlu, üzgün, korkmuş, şaşkın gibi. Burada sunulan çalışma, video dizilerinde yorgunluk analizi çalışmasının bir ön çalışması olarak yapılmaktadır. Bu nedenle göz öznitelik noktaları başlangıç noktası olarak alınmıştır. Çalışmada amaç, yetersiz ve kalitesiz verilerle dahi çalışan bir öznitelik takip algoritması geliştirmektir.

Literatürde var olan çalışmalarda öznitelik noktalarının takibi, ikiye ayrılmaktadır: Genel amaçlı olanlar, yüz öznitelik nokta takibi için olanlar. Genel amaçlı olan yöntemlerden nokta temelli olan yöntemler [7,8] yumuşak hareket, sınırlı hız ve görüntüde engel olmayacağı gibi yüz için uygun olmayan güçlü varsayımlarda bulunmaktadır. Çerçevesel arasındaki

parça ilintisine bakan yöntemler [9,10] ise imgelerdeki ışıklık değişimine ve duruş değişikliklerine duyarlılık göstermektedir. Optik akış yöntemleri [11] de video dizisi boyunca piksel değerlerinin değişmeyeceğini farz ettiği için ışıklık ve duruş değişikliklerine duyarlıdır.

Yüz için özgül yöntemler diğer yöntemlere göre daha verimli sonuçlar vermektedir. Gabor süzgeçleri [1,2] yüz özniteliklerinin takibinde kullanılmış bir yöntemdir. Etkin şekil modelleri ve etkin görünüm modelleri [12,13] kullanılarak özniteliklerin belli bir modele oturtulduğu çalışmalar da vardır. Ayrıca [4]'te öznitelik noktaları arasında parametrik olmayan yöntemler kullanılarak uzamsal ve zamansal bir bağ kurulmaktadır.

Genel olarak öznitelik noktalarının takibi piksel değerlerine bağlı zamansal bir modele oturtulmaktadır. Bunun sonucunda ışıklık değişimlerine duyarlı, duruş değişikliklerini göz ardı eden, noktalar arasındaki uzamsal bilgiyi kullanmayan yaklaşımlar ortaya çıkmıştır. Bu yaklaşımlar, gürültülü imgelerde, noktalarının önüne engel geldiği, noktaların kaybolduğu gibi durumlarda fiziksel olarak mantıksız olan kaymalara, takip işlemini güçleştirmeye, dolayısıyla kötü sonuçlara yol açmaktadır. Uzamsal bağlantının kurulduğu [4]'te ise yöntem parametrik olmayan kestirimlere dayandığı için hesapsal olarak yavaş çalışmaktadır.

Bu makalede öznitelik noktaları, zamansal ve uzamsal bilgilerin bir araya getirildiği bir yöntemle takip edilmektedir. Noktalar arasındaki uzamsal bağlantıyı kurabilmemiz son zamanlarda birçok yapay görme konularında uygulanan grafiksel modellere dayalı çerçevemiz sayesinde mümkün olmaktadır. Yaklaşımda parametrik kestirimler kullanıldığı için hesapsal yük azdır. Kurulan bu bağlantı ile gözden kaybolan bir öznitelik noktası uzamsal komşuları sayesinde gözlem yapılamasa bile takip edilebilmektedir. Video dizilerinden öznitelik noktalarının sezimi için Gabor süzgeçleri kullanılmaktadır. Süzgeçler imgelerdeki kenar bilgilerini algılayan; farklı duruş, dönme açılarına ve farklı boyuttaki özniteliklere duyarlı olacak şekilde kullanılmaktadır. Gerçek video verileri üzerinde yapılan sonuçlar göstermiştir ki, önüne engel gelerek kaybolan öznitelik noktaları uzamsal komşularından gelen bilgi kullanılarak başarıyla takip edilebilmektedir.

2. Önerilen Yöntem

2.1. Öznitelik Noktası Takip Problemi ve Ön İşleme

Bu makalede çözüm önerisi getirilen problem, 2-B gri video dizilerindeki gözün iki uç noktası olarak seçilen öznitelik

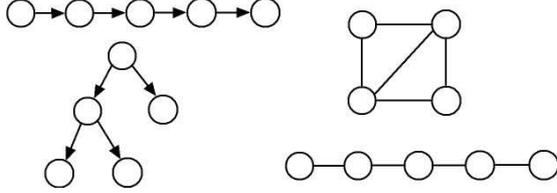
noktalarının dizi boyunca takip edilmesidir. Önerilen yöntemin gözlem verileri Gabor süzgeçlemesi gibi bir ön işleme sonucu elde edilmektedir.

Kullanılan süzgeç fonksiyonu [3]'deki gibi seçilmektedir. Daha sonra ardışık iki video çerçevesi [1]'e göre 6 farklı açı ve 4 farklı frekans bandı için elde edilen 24 adet süzgeçten geçirilmektedir. Süzgeçleme sonucu gabor uzayındaki karmaşık değişkenlerin genlik ve faz bilgileri [3]'deki benzerlik metriğine sokularak, şu anki noktanın bir sonraki çerçevede nerede olabileceği sezgisel olarak öngörülmektedir. Ön görülen bu nokta önerilen yöntemin gözlem verisidir.

2.2. Grafiksel Modeller

Grafiksel modeller, grafik teorisi ile olasılık teorisinin birbirinin içine geçerek oluşturduğu bir yapıdır. Bu yapıda, olasılık teorisi kullanılarak oluşturulan modeller, grafik teorisinin görsel gücü sayesinde hem daha rahat anlaşılmakta hem de büyük modellerdeki ilişkiler daha rahat kurulmaktadır.

Genel olarak G ile gösterilen bir grafik, düğüm kümesi V 'den ve kenar kümesi E 'den oluşmaktadır. Modeldeki $s \in V$ düğümünün komşuları $N(s) = \{t | (s,t) \in E\}$ olarak tanımlanmaktadır. Kenarların yönlü ya da yönsüz olmasına göre modeller ikiye ayrılmaktadır. Yönlü modeller nedensel ilişkinin olduğu modellerdir. Yönsüz modeller ise ilişkinin çift yönlü olduğu modellerdir. Bazı grafiksel modellere örnekler Şekil 1'de gösterilmektedir.



Şekil 1: Bazı örnek grafiksel modeller

Grafiksel modellerde her düğüm, saklı (x_s) ya da gözlemlenen (y_s) rastlantı değişkenini göstermektedir. Genel olarak $x = \{x_s | s \in V\}$ ve $y = \{y_s | s \in V\}$ ile tüm saklı ve gözlemlenen değişkenleri tanımlarsak, yönsüz bir modeldeki birleşik olasılık yoğunluk fonksiyonu $p(x,y)$ aşağıdaki şekilde bileşenlere ayrılabilir:

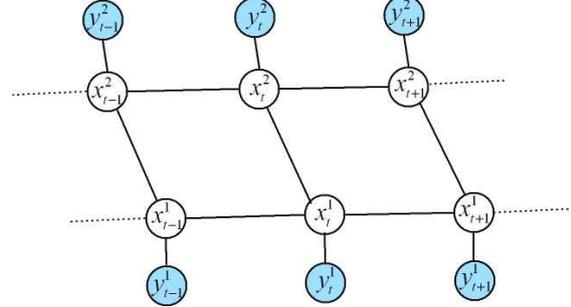
$$p(x, y) = \frac{1}{Z} \prod_{(s,t) \in E} \psi_{s,t}(x_s, x_t) \prod_{s \in V} \psi_s(x_s, y_s) \quad (1)$$

Burada $\psi_{s,t}(x_s, x_t)$ ile gösterilen fonksiyon saklı değişkenler arasındaki kenar potansiyel fonksiyonu göstermektedir. Diğer terim $\psi_s(x_s, y_s)$ ise gözlem modelinin fonksiyonudur.

Uygulanan yöntemde kullanılan grafiksel model Şekil 2'de gösterilmektedir. Modeldeki her saklı düğüm (x_s) 4 elemanlı bir rastlantı vektörüdür. İmgelerdeki kafa hareketinin 2-B olduğu düşünüldüğü için bu dört eleman sırasıyla; öznitelik noktalarının x eksenindeki koordinatı, y eksenindeki koordinatı, x eksenindeki hızı ve y eksenindeki hızıdır.

Gözlemlenen düğümler (y_s) ise 2 elemanlı olup, sadece x ve y koordinatlarına tekabül eden ölçümleri göstermektedir.

Buna göre (x_t^1) birinci öznitelik noktasının t anındaki saklı değişkenini, (y_{t+1}^2) ise ikinci öznitelik noktasının t+1 anındaki gözlemlenen değişkenini göstermektedir.



Şekil 2: Kullanılan grafiksel model

Buradaki kenar potansiyel fonksiyonları seçimi 2.3–2.5 bölümlerinde anlatılmaktadır.

2.3. Zamansal Model

Zamansal model öznitelik noktalarının zamansal olarak bir önceki değeri ile olan ilişkisini göstermektedir. Bu model aşağıda gösterilen geçiş modeline dayanmaktadır.

$$x_{t+1} = A \cdot x_t + w \quad w \sim N(0, Q) \quad (2)$$

Buradaki A geçiş matrisi, Q ise sıfır ortalama değerli normal dağılımlı gürültünün ortak değişim matrisidir. Bu değerler aşağıdaki gibi tanımlanmaktadır:

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad Q = \begin{pmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & \sigma_v^2 \end{pmatrix}$$

Sonuç olarak zamansal iki komşu arasındaki ilişki normal dağılımlı bir ilişkidir ve aşağıda verilen kenar potansiyel ile ifade edilebilir:

$$\psi_{t,t+1}(x_t, x_{t+1}) = \alpha \exp \left\{ -\frac{1}{2} [x_{t+1}^T \ x_t^T] \begin{bmatrix} Q^{-1} & -Q^{-1}A \\ -A^T Q^{-1} & A^T Q^{-1}A \end{bmatrix} \begin{bmatrix} x_{t+1}^T \\ x_t^T \end{bmatrix} \right\} \quad (3)$$

2.4. Uzamsal Model

Uzamsal model, öznitelik noktaları arasındaki uzamsal komşuluk ilişkisinin kurulduğu yerdir. Bu ilişki basit bir biçimde noktalar arasında olması beklenen uzamsal mesafe bilgisini kullanacak şekilde seçilmektedir. Gözün iki uç noktası arasındaki mesafe buna örnek olabilir. Bunlara dayanarak uzamsal ilişki aşağıdaki şekilde kurulmaktadır:

$$\psi_{1,2}(x_1, x_2) = \alpha \exp \{ [x_1 - (x_2 - \Delta x)]^T \Sigma^{-1} [x_1 - (x_2 - \Delta x)] \} \quad (4)$$

Buradaki Δx noktalar arasındaki mesafeyi göstermektedir. Ortak değişim matrisi aşağıdaki şekilde tanımlanmaktadır:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & \sigma_v^2 \end{pmatrix}$$

2.5. Gözlem Modeli

Gözlem verilerinin, y , video dizilerinden nasıl elde edildiği 2.1 bölümünde anlatılmaktadır.

Gözlem modeli, saklı rastlantı değişkeni x_s ile gürültülü yerel gözlem y_s arasındaki ilişkiyi kurmaktadır. Bu model şu şekilde düşünülmüştür:

$$y_t = C \cdot x_t + v \quad v \sim N(0, R) \quad (5)$$

Buradaki C gözlem matrisini, R ise sıfır ortalama değerli normal dağılımlı gürültünün ortak değişinti matrisini göstermektedir. Bu değişkenler aşağıdaki gibi tanımlanmaktadır:

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad R = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

Bunun sonucu olarak x_s ile y_s arasındaki ilişki normal dağılımlı bir ilişkidir ve aşağıdaki şekilde tanımlanabilir:

$$\psi_s(x_s, y_s) = N(C \cdot x_s, R) = p(y_s | x_s) \quad (6)$$

2.6. Döngülü İnanç Yayılım Algoritması

Yapay görme ve görüntü işleme uygulamalarında tipik olarak $p(x_s | y)$ koşullu olasılık yoğunluk fonksiyonu bulunmaya çalışılır. Döngüsüz grafiksel modellerde bu koşullu olasılık doğrudan *inanç yayılımı* olarak da bilinen yerel mesaj-geçirme algoritması kullanılarak hesaplanabilmektedir. Bu algoritma zincir yapıları modellerde tipik olarak Kalman ya da Parçacık süzgeçlemeye, düzeltirmeye denk gelmektedir. Döngülü grafikler için Pearl [6] göstermiştir ki *inanç yayılımı* algoritması işe yaramaktadır. Bu algoritma şu şekildedir: Her $t \in V$ düğümü komşusu olan $s \in N(t)$ düğümüne göndermek üzere bir $m_{ts}(x_s)$ mesajı oluşturur:

$$m_{ts}(x_s) = \alpha \int_{x_t} \psi_{s,t}(x_s, x_t) \psi_t(x_t, y_t) \times \prod_{u \in N(t) \setminus s} m_{tu}(x_t) dx_t \quad (7)$$

Her düğüm kendisine gelen bu mesajlar ile kendi gözlemini birleştirerek kendi koşullu olasılık fonksiyonunu elde etmektedir:

$$p(x_s | y) = \alpha \psi_s(x_s, y_s) \prod_{t \in N(s)} m_{ts}(x_s) \quad (8)$$

Kurulan modeldeki tüm ilişkiler Gauss (Normal) dağılımlı seçildiği için algoritmanın yukarıda bahsedilen iki adımı ortalama değer ve ortak değişinti matrislerinin güncellenmesi üzerine kurulmaktadır. Bu nedenle algoritma parametrik olmayan yöntemlere göre daha hızlı çalışmaktadır. Saklı rastlantı vektörü x ve y eksenindeki koordinat ve hızlar olarak tanımlandığından normal dağılımların ortalama değerleri aynı şekilde x ve y eksenindeki koordinat ve hızların kestirim değerleridir. Her güncelleme adımı yeni bir gözlem yapıldığında o ana kadar gelen gözlemler kullanılarak

yapıldığı için de yöntem tipik bir süzgeçlemedir. Güncellenme denklemleri tipik olarak [5]'deki gibi seçilerek kullanılmıştır.

3. Deneysel Sonuçlar

Önerilen yöntem, yöntemin başarımını göstermek amacıyla basit olarak gözdeki iki öznelik noktası üzerinde denenmiştir. Denemenin yapıldığı video dizi kümesi kafanın sağa ve sola ötelenmesini, bir ve birkaç noktanın bir engelle kaybolmasını içermektedir. Zamansal, uzamsal ve gözlem modelindeki ortak değişinti matris parametreleri bu dizilere uygun olarak seçilmiştir. Takip sonucu çıkan imgeler Şekil 3'de gösterilmektedir. Daha önceki çalışmalarda gibi sadece zamansal ilişkiye bağlı olan model ile yapılan takip sonucu çıkan imgeler ise Şekil 4'de gösterilmektedir. Her şekilde gösterilen yeşil noktalar algoritmanın o ana kadar gözlemlenen verileri kullanarak kestirdiği çıktılardır.

Şekil 3'de görüldüğü gibi önerilen algoritma kafanın sağa-sola ötelenmesi, bir noktanın engelle kaybolması gibi durumlarda başarıyla takibi sürdürmektedir. Diğer yandan sadece zamansal ilişki üzerine kurulu model ile yapılan takipte noktalar farklı yerlere kaymakta, takip düzgün sürdürülememektedir.

Gözün önünden parmak geçirilerek noktaların engellendiği durumlarda, bir engel algılayıcının var olduğu farz edilerek engelin olduğu bölümler gözlem çıkarım işleminden atılmaktadır. Burada bir mükemmel çalışan, bir de çok iyi çalışmayan, engel olmadığı halde engel var gibi gösterip veri kaybına yol açan iki tip engel algılayıcının olduğu düşünülmüştür. Bu iki algılayıcı ile de çalıştırılan algoritma sonuçlarından gözle görülecek bir fark gözükmemiştir. Buradan da anlaşılacağı gibi veri kaybı olduğu halde takibi doğru gerçekleştiren algoritma gürbüz bir takip algoritmasıdır.

4. Sonuç ve İleriki Çalışmalar

Bu makalede, gürbüz çalışan bir göz öznelik nokta takip algoritması geliştirilmiştir. Algoritmadaki en büyük avantaj noktalar arasındaki sadece zamansal ilişkiyi değil uzamsal ilişkiyi de göz önüne almasıdır. Bu sayede, herhangi bir noktanın kaybolduğu, engellendiği durumlarda komşu noktalar aracılığıyla takip devam edebilmektedir. Algoritmanın diğer bir avantajı da parametrik kestirim yöntemleri kullandığı için hesapsal olarak parametrik olmayana nazaran daha hızlıdır. Sadece iki göz öznelik noktası için uygulanan bu yöntem ileriki çalışmalarda uygun model yapısı ve parametre seçimi yapılarak tüm yüz öznelik noktaları için uygulanacaktır.

5. Kaynakça

- [1] H Gu, Q Ji, "Information extraction from image sequences of real-world facial expressions" Mach. Vis. Appl. 16(2):05
- [2] Q Ji, X Yang "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance" Real-Time Imaging 8(5): 357-377 (2002)
- [3] D. S. Bolme, "Elastic Bunch Graph Matching", Ms, 2003
- [4] C Su, L Huang, "Spatio-Temporal Graphical-Model-Based Multiple Facial Feature Tracking", Jnl on Appld Signal Proc. 2005:13

- [5] E B. Sudderth, "Embedded Trees: Estimation of Gaussian Processes on Graphs with Cycles", Master Thesis, 2002
- [6] J. Pearl. "Probabilistic Reasoning in Intelligent Systems" Morgan Kaufman, San Mateo, 1988.
- [7] Chetverikov D, Verestoy J, "Tracking feature points: a new algorithm", Proceedings of the international conference on pattern recognition, pp 1436–1438 1998
- [8] Rangarajan K, Shah M, "Establishing motion correspondence" CVGIP:Image Underst. 54:56–73 1991
- [9] Bretzner L, Lindeberg T, "Feature tracking with automatic selection of spatial scales", Comp. Vis Image Under. 71(3):98'

- [10] Shapiro L, Wang H, Brady J, "A matching and tracking strategy for independently-moving, non-rigid objects", Proceed. of the 3rd British mach visi conf., 306–315 1992
- [11] Meyer F, Bouthemy P, "Region-based tracking using affine motion models in long image sequences" CVGIP Image Understand 60:119–140 1994
- [12] Cootes TF, Edwards GJ, Taylor CJ, "Active appearance models", Proceedings of the European conference on computer vision 2:484–498 1998
- [13] J Xiao, S Baker, I Matthews, T Kanade, "Real-Time Combined 2D+3D Active Appearance Models" CVPR04



(a)



(b)



(c)

Şekil 3: (a) Öteleme dizisi (b) Yatay engel dizisi (c) Dikey engel dizisi uygulanan algoritma sonuçları



(a)



(b)



(c)

Şekil 4: (a) Öteleme dizisi (b) Yatay engel dizisi (c) Dikey engel dizisi sadece zamansal ilişkiye bağlı algoritma sonuçları