

Saçılım Analizi Temelli bir Göz Hareketi Olay Belirleme Algoritması (Türkçe)

Density Based Eye Movement Event Detection Algorithm (In English)

Ozan Deniz, Mehmetcan Fal, Cengiz Acartürk
Bilişsel Bilimler Programı, Enformatik Enstitüsü
Orta Doğu Teknik Üniversitesi
06800 Ankara, Türkiye
{ozan.deniz, mehmetcan.fal, acarturk}@metu.edu.tr

Özetçe—Göz izleme cihazlarının yaygınlaşması ile ham veriden göz hareketi çıkarsayan algoritmalar önem kazanmıştır. Günümüzde yaygın olarak kullanılan algoritmaların ortak özelliği, veri noktalarının hız ve konum bilgilerini ayrı ayrı ele almasıdır. Bu çalışmada mevcut algoritmaların güçlü ve zayıf yanları incelenmiş, göz hareketlerinin oluşturduğu dağılımı esas alan yeni bir yaklaşım geliştirilmiştir.

Anahtar Kelimeler—Göz izleme, göz hareketleri., göz hareketi algoritmaları, insan bilgisayar etkileşimi

Abstract—Extracting eye movement events from raw data has become important upon the introduction of affordable eye trackers. The common approach in current algorithms is that they take into account velocity and position information separately. In this paper, we analyze the pros and cons of the current algorithms and introduce a novel algorithm based on the distribution of raw eye movement data.

Keywords—Eye tracking, eye movements., eye movement algorithms, human computer interaction

I. GİRİŞ: GÖZ İZLEME TEKNOLOJİLERİNDE MEVCUT DURUM VE YÖNELİMLER

Göz izleme (eye tracking) analizi son 20 yıldır okuma, görsel algı, çevre algısı gibi temel bilişsel psikoloji çalışmalarından eğitim araştırmalarına ve kullanılabilirlik uygulamalarına kadar birçok alanda kullanılan analiz yöntemlerinden birisidir [1]. Bakışın (gaze) gözlenebilir ve sayısal olarak kaydedilebilir olması, dikkat ve öğrenme gibi içsel süreçlerle ilişkisi ve gözün hareket karakteristiklerinin çok sayıda ölçüm

parametresinin elde edilebilmesine olanak tanınması göz hareketleri (eye movement) analizlerinin ilgili araştırma alanlarındaki önemini korumaya devam edeceğine işaret etmektedir. Türkiye’de göz hareketleri izleme çalışmaları son 10 yıldır yürütülmektedir [2,3]. Son yıllarda azalan alıcı cihaz maliyetleri nedeniyle göz izleme analizinin insan-bilgisayar etkileşimi alanında, özellikle çok-kipli (multi-modal) etkileşim sistemleri tasarımında kullanılacağı öngörülebilir [11,12].

Günümüzde çoğunlukla masaüstü bilgisayarların monitörlerine ya da dizüstü bilgisayara bağlanarak kullanılan alıcı cihazlar gözün anlık konumunu *nitelikli ham veri*¹ olarak sağlamaktadır. Bu veriden yola çıkarak sabitleme (fixation), sıçrama (saccade) gibi temel göz hareketlerini belirleyen algoritmalara ihtiyaç duyulmaktadır. Bu çalışmanın amacı ham göz izleme verisini göz hareketi olaylarına dönüştüren algoritmaları tanıtmak ve araştırmacılar tarafından geliştirilen, saçılım analizi temelli bir göz hareketi olay belirleme algoritmasını araştırmacılara sunmaktır. Sonraki bölümde ham veriden elde edilebilecek göz hareketi olay türleri tanıtılmaktadır.

II. HAM VERİDEN ELDE EDİLEBİLECEK GÖZ HAREKETİ OLAY TÜRLERİ

Göz hareketi olayları *sabitleme* ve *izleme* olarak iki temel fonksiyon çerçevesinde sınıflandırılabilirler. Sabitleme ve sıçrama günlük yaşamda en sık gözlenen göz hareketi olaylarıdır. Sabitleme, gözün belirli bir konumda belirli bir süre sabit durmasını ifade eder.

¹ “Nitelikli ham veri” terimi ile göz izleme alıcısı (eye tracker sensor) tarafından gönderilen kızılötesi ışınların gözün kornea ve retinasından gelen yansımalarının işlenmesi ile elde edilen ve SDK kütüphaneleri aracılığı ile geliştirici kullanıcılara sunulan işlenmiş veri kastedilmektedir.

Okuma sırasında 200-300 ms civarında gerçekleşir. Diğer görsel durumlarda sabitleme süreleri bağlama göre değişiklik gösterir. Sabitlemeler arasındaki *sıçramalar* da çevresel bağlama göre farklı özellikler sergiler. Örneğin okumada metnin dilsel özelliklerine bağlı olarak değişen sabitleme süreleri, konumları ve yatay yönde ilerleyen sıçramalar gözlenir. İzleme (pursuit) ise nesnenin ya da kafanın hareketli olduğu durumlarda takip edilmesine olanak sağlayan göz hareketlerini ifade eder [4].

Klinik olmayan göz izleme analizi çalışmalarında temel olarak belirlenmeye çalışılan göz hareketi sabitlemedir. Bu çalışmada sabitleme tespiti üzerinde durulmaktadır. Aşağıdaki bölümde nitelikli ham veriden sabitleme tespit eden göz hareketi algoritmaları tanıtılmaktadır.

III. GÖZ HAREKETİ OLAY BELİRLEME ALGORİTMALARI

Günümüzde ham veriyi kullanarak göz hareketi olaylarını belirleyen, özellikle sabitleme ve sıçrama tespiti yapan, veri noktalarını sabitleme ya da sıçrama olarak iki ana sınıftan birisine ait olarak etiketleyen algoritmalar, temel olarak veri noktalarının (dolayısıyla bakışın ve gözün) konumuna ve konumun değişim özelliklerine odaklanmaktadır. Mevcut algoritmalar üç ana grup altında toplanabilir.

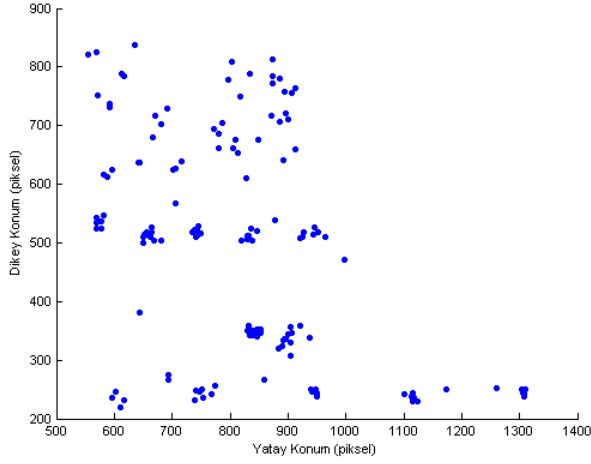
- Gözün hızını temel alan algoritmalar
- Bakışın konumuna bağlı alan-tabanlı algoritmalar
- Verinin saçılım özelliklerinin incelenmesine dayanan algoritmalar

Gözün hızını temel alan algoritmalar örnek olarak I-VT (Velocity-Threshold Fixation Identification), I-HMM (Hidden Markov Model Fixation Identification) ve I-KF (Kalman Filter Fixation Identification) algoritmaları; bakışın konumunu temel alan alan-tabanlı algoritmalar örnek olarak I-AOI (Area-of-Interest Fixation Identification) algoritması; verinin saçılım özelliklerinin incelenmesine dayanan algoritmalar örnek olarak I-DT (Dispersion-Threshold Fixation Identification) ve I-MST (Minimum Spanning Trees Fixation Identification) algoritmaları verilebilir [5,6].

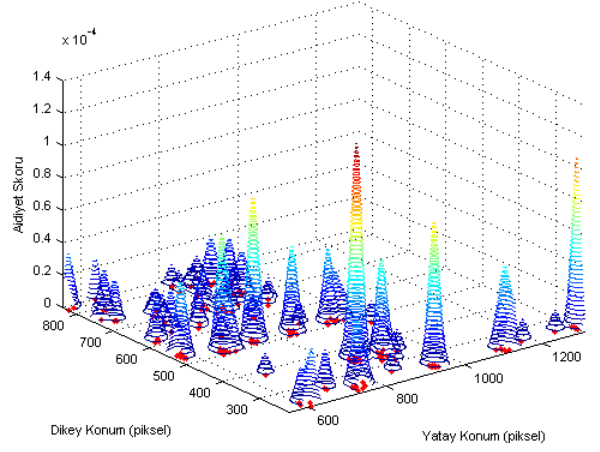
Tüm göz hareketi olay belirleme algoritmaları girdi olarak örnekleme noktasındaki veri noktalarının yatay ve dikey koordinatlarını kartezyen ya da dairesel referans koordinat sistemine göre alır. *Dairesel referans koordinat sistemini* kullanan algoritmalar ekranın genişlik ve yükseklik değeri ile uyarının göz pozisyonuna olan uzaklığı dikkate alınmalıdır [6]. Göz izleme alıcısının ardışık kayıtları arasındaki zaman farkı da girdi olarak alınır. Varsayılan koordinat sistemine bağlı olarak x_i , y_i , t formatında bir veri seti oluşturulur.

Ham veriden sabitleme ve sıçrama tespiti yapabilmek amacıyla en sık kullanılan algoritmalarından birisi I-VT'dir

(Velocity-Threshold Fixation Identification). Ardışık veri noktaları arasındaki mesafe hesaplanarak gözün hareket hızını bulan algoritma, belli bir eşik değerinin altında kalan hızları *sabitleme*, diğerlerini *sıçrama* olarak sınıflandırır. Sonraki aşamada algoritma sıçrama olarak sınıflanan verileri işlem dışı bırakır ve ardışık sabitleme noktalarını gruplar. Her bir grup içinde yer alan sabitleme veri noktaları aralarındaki mesafelerin orta noktasından, onu karşılayan noktaya ulaşan doğruların kesişmesiyle elde edilen ortanca (centroid) nokta, odaklanma noktası olarak kabul edilir. I-HMM (Hidden Markov Model Fixation Identification) algoritması da prensip olarak I-VT ile aynı mantığı uygulamaktadır. I-KF (Kalman Filter Fixation Identification) hız temelli göz hareketi olay belirleme algoritmalarından bir diğeridir. Bu algoritma oluşabilecek en yüksek göz hızını kullanarak veri setini örneklemin gürültü olduğu varsayımı ile inceler ve veri setini oluşturan göz hızı hakkında bir tahminde bulunur. Bu tahmini değer ve veri setinden elde edilen göz hızı değeri Chi-square testi için kullanılır. Böylece veri içerisindeki sıçrama ve sabitleme olayları birbirinden ayrılır [6]. Veri noktalarının konum analizine dayalı alan-tabanlı algoritmalarından I-AOI (Area of Interest Fixation Identification) algoritması, hız-tabanlı algoritmalarından farklı olarak belirli bir alandaki veri noktalarını *sabitleme*, dışındaki noktaları ise *sıçrama* olayı olarak sınıflandırır. Bu algoritmanın kullanımındaki dezavantajlardan birisi tespit edilen alan içerisinde farklı zamanlarda oluşan birden fazla sıçrama ve sabitleme davranışını ayırt edememesidir [6]. Verinin saçılım özelliklerinin incelenmesine dayanan algoritmalarından birisi olan I-DT'de (Dispersion-Threshold Fixation Identification) diğer algoritmalarından farklı olarak kullanıcı ardışık noktaların birbirleri arasındaki en fazla uzaklığı bir eşik değeri olarak belirler. Bu uzaklığın altında kalan tüm noktaların bir küme olduğu varsayımı ile kümenin içerisindeki ortanca bulunur ve odaklanma noktası olarak işaretlenir. Yeni küme oluşturmak için takip eden ardışık veriler tüm veri bitene kadar aynı işleme tabi tutulur [6]. Saçılım-analizi tabanlı bir diğer algoritma olan I-MST (Minimum Spanning Tree Fixation Identification) veri seti içerisinde birbirlerine bağlı noktalar (Spanning Tree) arasındaki en küçük Öklid uzaklığını hesaplayarak, veri seti içerisinde görece küçük uzaklıkları barındıran birbirlerine bağlı noktalar kümesini sabitleme olarak işaretler ve görece büyük uzaklıkları da sıçrama olarak kabul eder. Öklid uzaklıklarını barındıran bu kümeleri oluşturmak için kullanılan eşik değeri en büyük sıçramanın oluşabileceği zamansal bir aralıktır. Bu değer, birbirine bağlı bu kümeleri oluşturmak için gerekli olan periyodu belirler (Window Size). Saçılım-analizi tabanlı algoritmaların güçlü yanı, gürültüden dolayı



Şekil 1. Katılımcının Bakış Noktaları



Şekil 2. Çok Değişkenli Çekirdek Yoğunluk Tahmini Yöntemi Sonrasında Katılımcının Bakış Noktaları

oluşan veri kaybından etkilenmeden doğru odaklanma noktalarını tanımlayabilmesidir [6]. Aşağıdaki bölümde mevcut çalışma kapsamında geliştirilen saçılım analizi temelli bir göz hareketi olay belirleme algoritması tanıtılmaktadır.

IV. SAÇILIM ANALİZİ TEMELLİ BİR GÖZ HAREKETİ OLAY BELİRLEME ALGORİTMASI

Algoritma, sabitleme olayının oluşturduğu saçılımı temel almaktadır. Saçılımın merkez noktası, dağılımın tepe noktasına karşılık gelmektedir. Algoritma saçılımda bulunan noktaların merkezden uzaklığına göre, aidiyet skoru hesaplanmaktadır. Aidiyet skorları *kartezyen koordinat sistemi* temel alınarak hesaplanmaktadır.

Sabitleme ve sıçrama olayları arasındaki fark, noktaların aidiyet skorlarının oluşturduğu yoğunluk bilgisine göre belirlenmektedir. Düşük aidiyet skorlu noktalar *sıçrama* olayı, yüksek aidiyet skorlu olan noktalar ise *sabitleme* olayı olarak sınıflandırılmaktadır.

Algoritma için ham göz verisi toplanırken kullanılan cihazın modeli Tobii X2-60 ve verinin kalitesi %90 olarak belirlenmiş, katılımcılara harflerden oluşan bir ekran gösterilmiş ve katılımcılar ekrana bakarken, uyarın altında göz verileri toplanmıştır. Veri yatay ve dikey eksenlerdeki bakış noktalarından oluşmaktadır ve içerisinde gerileme (regression) olayı bulunmamaktadır. Deneklerin sabitleme olayını gerçekleştirdiği uyarılara, tekrar geri dönemeyecekleri şekilde uyarın ayarlanmıştır.

Algoritma üç adımdan oluşmaktadır. Birinci adımda aidiyet skorları çok değişkenli çekirdek yoğunluk tahmini [7] (multivariate kernel density estimation) yöntemi ile hesaplanır. Her bir noktanın aidiyet skorunun, yoğunluk dağılım fonksiyonuna yaptığı katkı Gaussian çekirdeği kullanılarak yumuşatılır (ing. smoothing) [8]. Çok değişkenli çekirdek yoğunluk tahmini yöntemindeki en

önemli problem bant genişliği parametresinin belirlenmesidir [8]. Bant genişliği yüksek belirlenir ise, her bir bakış noktasının aidiyet skorunun, saçılım yoğunluk fonksiyonuna yaptığı katkı fazla yumuşatılmış olur (oversmoothing). Bant genişliği düşük seçilir ise saçılım yoğunluk grafiğinde sivrilikler (spikes) açığa çıkar. Bant genişliği, MISE (ortalama bütünleşik karesel yanılğı) değeri kullanılarak hesaplanmalıdır. MISE değerinin hesaplanması matematiksel olarak kompleks olduğu için algoritmada matematiksel olarak MISE yerine geçebilecek olan AMISE (asimtotik ortalama bütünleşik karesel yanılğı) değeri kullanılmıştır. Algoritma AMISE değerini kullanarak bant genişliğini hesapladıktan sonra bu değer ile oluşturulan Gaussian çekirdeğini veriye uygulayarak, noktaların aidiyet skorlarını hesaplar. Çok değişkenli çekirdek yoğunluk tahmini yöntemi veriyi eşit ve ikinin kuvveti olacak şekilde yeni bir düzlem üzerine taşır. Örneğin 114 adet nokta, uygulanan yöntem sonucunda 128x128 boyutunda bir matris üzerine oturtulur. Matrisin her bir elemanı, koordinat sisteminde karşılık gelen nokta ya da noktaların yoğunluğunu temsil etmektedir. Şekil 1’de tek bir katılımcıya ait bakış noktaları ham veri olarak gösterilmiş, Şekil 2’de kaydedilen veri noktalarının matris üzerinde modellenmiş hali gösterilmiştir.

İkinci adımda *k-means kümeleme algoritması* kullanılarak aidiyet skorları, yoğunluk düzeyleri olarak kümelendirilir. *K-means* öğreticiye dayalı olmayan (unsupervised) bir kümeleme algoritmasıdır; en büyük dezavantajı veride kaç adet kümenin bulunduğu önceden belirlenmesi gerekliliğidir. *K-means*, noktaların birbirine yakınlığını baz alır ve belirlenen küme sayısına erişilene kadar kümeleme işlemini sürdürür. Algoritmada küme sayısı belirlenirken, kümelemenin yol açtığı

bozulmayı en aza indirmeyi amaçlayan metrik kullanılmaktadır [9]. *K-means kümeleme algoritması* için kullandığımız metriğe göre en iyi k değerleri 2 ve 8 olarak belirlenmiş, algoritmada k değeri 8 olarak seçilmiştir. Veri yüksek sayıda kümelendirilirse, göz olayları daha kesin bir şekilde sınıflandırılabilir. Bu yüzden k değeri olarak büyük olan değer seçilmiştir. *K-means* ile noktaların aidiyet skorları kümelendikten sonra, yoğunluk düzeyleri belirlenir. Algoritmaya göre veri içerisinde herhangi bir kümeleme 8 farklı yoğunluk düzeyinden oluşmaktadır.

Üçüncü adımda sınır belirleme algoritması kullanarak her bir yoğunluk düzeyi içerisindeki kümelenecekler birbirinden ayrılır. Bu noktada alternatif bir çözüm *k-means* ya da benzer bir kümeleme algoritması kullanılmaktadır. Ancak *k-means* kullanılan durumda, her bir düzey için k değerinin tekrardan hesaplanması gerekmektedir. Düzeyler içerisindeki kümelenecekler birbirinden fiziksel olarak bağımsız olduğundan, denetimsiz bir sınır belirleme algoritması kullanmak ve bu kümelenecekleri birbirinden ayırmak *k-means kümeleme algoritmasına* kıyasla matematiksel olarak daha basit bir çözüm sunmaktadır.

Veride bir düzeye ait olan kümelenecekler ele alınırken, diğer düzeylere ait olan noktalar sınır belirleme algoritmasına dahil edilmemektedir. İşlem en yüksek aidiyet noktalarının bulunduğu düzeyden en düşük aidiyet noktalarının bulunduğu düzeye doğru ilerlemektedir. Her bir düzeyde oluşan kümelenecekler *Moore-neighborhood sınır belirleme algoritması* [10] kullanılarak birbirinden ayrılır. Küme içerisindeki boşluklar bize sabitleme ya da sıçrama olayı hakkında bilgi vermektedir. İçerisinde boşluk bulduran küme sıçrama olayıdır. Çünkü bir olayın sabitleme olayı olabilmesi için göz noktalarının belirli bir yoğunluk düzeyinde homojen olarak gruplanması gerekmektedir. Son aşamada kümelenecekler içerisindeki boşlukları modellemek için içbükey kabuk (ing. convex hull) algoritması kullanılmıştır.

V. SONUÇ

Mevcut çalışma kapsamında geliştirilen göz hareketleri belirleme algoritması, veri saçılımının dağılımını esas aldığı için veriyi kümülatif olarak ele almaktadır. Veriden *aidiyet skoru* adında yeni bir özellik ortaya çıkarmıştır. Aidiyet skoru, saçılım içerisinde bulunan noktaların birbirleri ile ilişkisini göstermekte ve göz hareketi olaylarının belirlenmesinde önemli rol oynamaktadır. Algoritmanın avantajları, ekipmandan kaynaklanan hataya karşı daha dirençli bir yapıya sahip olması ve veri işleme sırasında veri kaybını azaltmaya yönelik tasarımıdır. Gerçekleştirdiğimiz kıyaslama

analizleri algoritmanın literatürdeki algoritmalara yakın sabitleme sayıları döndürdüğünü göstermiştir (örn. bu çalışmada kullanılan data seti için; Saçılım Algoritması: 12, I-VT: 13, I-DT: 8, I-HMM: 13, I-KF: 10, I-MST: 14).

BİLGİLENDİRME

Bu çalışma 113K723 Okumada Bilişsel Süreçlerin İncelenmesi: Göz Hareketleri Kontrol Modellemesi için Türkçe Okuma Örüntüleri Derlemi başlıklı TÜBİTAK projesi ve 610986 IRIS: Towards Natural Interaction and Communication başlıklı EU FP7 projesi tarafından kısmi olarak desteklenmiştir.

KAYNAKÇA

- [1] Rayner, K., "Eye movements in reading and information processing: 20 years of research." *Psychological Bulletin*, 124(3), 372-422. 1998
- [2] Acartürk, C., & Çağiltay, K. (2006). İnsan Bilgisayar Etkileşimi ve ODTÜ'de yürütülen çalışmalar. 8. Akademik Bilişim Konferansı Bildirileri (pp. 335-340). Pamukkale Üniversitesi, Denizli.
- [3] Acartürk, C. (2004). Üniversite web sayfalarında kullanılabilirliğin önemi ve kullanılabilirlik testleri. 6. Akademik Bilişim Konferansı Bildirileri. Karadeniz Teknik Üniversitesi, Trabzon.
- [4] Henderson, J. M. (2012). Introduction to "Computational Approaches to Reading and Scene Perception". *Visual Cognition*, 20(4-5), 357-359.
- [5] Salvucci, D. D., Goldberg, J. H., "Identifying Fixations and Saccades in Eye-Tracking Protocols", *Eye Tracking Research and Applications Symposium* (pp. 71-78), 2000
- [6] Oleg V. Komogortsev, Denise V. Gobert, Sampath Jayarathna, Do Hyong Koh, and Sandeep M. Gowda, "Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors", *IEEE Trans. Biomed. Engineering* 57(11): 2635-2645, 2010
- [7] Botev, Z.I.; Grotowski, J.F.; Kroese, D.P., "Kernel density estimation via diffusion", *Ann. Statist. Volume 38, Number 5*, 2916-2957, 2010
- [8] Wand, Matthew P. and Jones, M. Chris "Kernel Smoothing", 1995
- [9] D T Pham, S S Dimov, and C D Nguyen, "Selection of K in K-means clustering", *Proc. IMechE Vol. 219 Part C: J. Mechanical Engineering Science*, 2004
- [10] P.Rajashekar Reddy, V.Amarnadh, Mekala Bhaskar, "Evaluation of Stopping Criterion in Contour Tracing Algorithms", *International Journal of Computer Science and Information Technologies*, Vol. 3, 3888-3894, 2012
- [11] Acartürk, C., Habel, C., Çağiltay, K., & Alaçam, Ö., "Multimodal comprehension of language and graphics: Graphs with and without annotations.", *Journal of Eye Movement Research*, 1(3):2, 1-15, 2008
- [12] Acartürk, C., & Habel, C., "Eye tracking in multimodal comprehension of graphs." In R. Cox & J. P. S. Diego (Eds.), *In Proceedings of the Diagrams Workshop on Technology-Enhanced Diagrams Research*. Vol. 887, pp. 11-25, Kent, UK: CEUR,2012