

Dinamik Zaman Bükme Metodu Kullanarak İvmeölçer Tabanlı El Yazısı Karakter Tanıma Accelerometer Based Handwritten Character Recognition Using Dynamic Time Warping Method

Esra Tunçer, Mehmet Zübeyir Ünlü

Elektrik Elektronik Mühendisliği Bölümü, İzmir Yüksek Teknoloji Enstitüsü, İzmir, Türkiye

{etuncer, zubeyirunlu}@iyte.edu.tr

Özetçe—Bu çalışmada, ivmeölçer kullanılarak el yazısı ile yazılan karakterlerin tanınması yapılmıştır. Karakter tanıma çalışmalarında genellikle kullanılan görüntü işleme teknikleri yerine, bu projede yazıyı yazan kişinin yazma hareketlerinden elde edilen veriler kullanılmıştır. Kişinin yazıyı yazma hareketlerini elde edebilmek için 3 eksenli ivmeölçer kullanılmış ve buradan elde edilen verilerle karakterler Dinamik Zaman Bükme yöntemiyle tanınmıştır. İvmeölçer ile elde edilen veriler genellikle gürültülü veriler olduğundan verilere tanıma işleminden önce filtreleme, bölütleme ve normalizasyon gibi ön işleme teknikleri uygulanmıştır. Yapılan deneysel çalışmalarda %98,08'lik doğru tanıma oranına ulaşılmıştır.

Anahtar Kelimeler — el yazısı karakter tanıma; 3 eksenli ivmeölçer; dinamik zaman bükme.

Abstract—In this work, a handwritten character recognition method is performed by using accelerometer. Usually, in character recognition studies, image processing methods are used. However in this work data obtained from writing movements of writers are used. To obtain the movements of writing process 3-axis accelerometer is used and Dynamic Time Warping method is applied for recognition. Accelerometer data are generally noisy so before applying Dynamic Time Warping they need to be filtered. In preprocessing part of our work, filtering, segmentation and normalization steps are applied. In experimental work, recognition accuracy rate of %98,08 is reached.

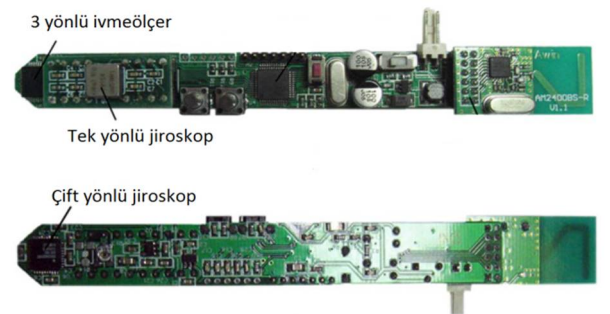
Keywords —handwritten character recognition; 3 axis accelerometer; dynamic time warping.

I. GİRİŞ

Günümüzde karakter tanıma en çok araştırılan ve popüler olan konulardan birisidir. Bununla beraber

bilgisayar tanımlı karakterleri tanıma hâlihazırda çok yüksek başarılarla ulaştığı için el yazısı tanıma ya da el yazısından yazan kişiyi tanıma konuları daha popüler olmaya başlamıştır. Çoğunlukla karakter tanıma işlemleri karakterin morfolojik yapısını kullanarak yapılmaktadır. Ancak bu şekilde bir tanımayla yazıyı yazan kişinin yazıyı yazma hareketleri bilinemediğinden tanımda bazı sorunlarla karşılaşılabilir. Özellikle kullanıcı tanımlı olmayan uygulamalarda bu sorun ortaya çıkmaktadır. Yazan kişinin yazma hareketlerini algılayabilmek için yazım esnasında el hareketlerini kamera ile kaydetme ya da ivme, açısal hız, kuvvet vb. değerleri ölçme gibi farklı yöntemler sıklıkla kullanılmaktadır.

Sung-Jung Cho ve Jong Koo Oh yaptığı çalışmada *Magic Wand* adını verdikleri cihazda kullanıcının el hareketlerinin ivmesini ve açısal hızını ölçen sensörler kullanmışlardır [1]. Benzer bir çalışma olan Jeen-Shing Wang ve Yu-Liang Hsu'nun çalışmasında *IMUPEN* adı verilen cihazda ise 1 adet 3 eksenli ivmeölçer, 2 adet jiroskop kullanılmıştır. Kullanılan jiroskoplardan bir tanesi tek eksenli jiroskop, diğeri 2 eksenli jiroskoptur [2].



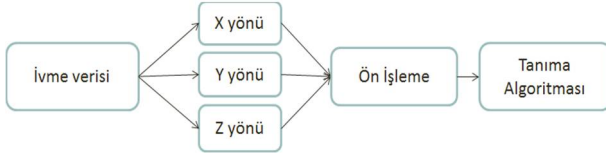
Şekil 1. IMUPEN [2]

Her ne kadar, bu tür çalışmalarda jiroskoplar çok kullanılıyor olsa da çok sayıda sinüs, kosinüs ve koordinat dönüştürme işlemleri yapılmalıdır. Bu işlemler işlemciye ağır hesaplama yükleri getirmektedir. Bu yükten kaçınmak için hesaplama, düşük maliyetli olmalı ve ağır hesaplama yükünden kaçınmalıdır. Yani eğer yapılmak istenen bir işlem sadece ivmeölçer ile yapılabiliyorsa o zaman jiroskop kullanılmamalıdır [3].

Tanım işleminden önce elde edilen veriler bir ön işlemeden geçirilmelidir. İvmeölçerler ile elde edilen ivme verileri genellikle yüksek frekanslı gürültüye sahip verilerdir. Bu yüksek frekanslı gürültüyü yok etmek için alçak geçiren filtre kullanılmalıdır. Literatürdeki çalışmalarda ivmeölçer verilerini filtrelemek için en yaygın olarak kullanılan filtre yürüyen ortalama filtredir. Ön işlemeden geçirilen veriler daha sonra tanıma yöntemleriyle incelenir. İvmeölçerlerle elde edilen verilerin zaman serileri olduğu göz önüne alınırsa buna uygun bir tanıma yöntemi seçilmelidir. Zaman serileri ile ifade edilen hareket aynı hareket bile olsa farklı tekrarlarla birebir aynı sinyali elde etmek neredeyse imkânsızdır. Aynı zaman diliminde aynı hareket yapılamayacağı gibi sinyalin zaman uzunluğu açısından da eşit sinyaller elde edilemez. Bu problemi çözmek için karşılaştırma ile tanıma yapılan çalışmalarda sıklıkla Dinamik Zaman Bükme (DZB) – Dynamic Time Warping (DTW) yöntemi kullanılmaktadır.

II. ÖN İŞLEME

Sensörden elde edilen 3 eksenli ivme verileri (X, Y, Z) öncelikle ayrı ayrı ön işlemeden geçirilmeli daha sonra tanıma algoritması uygulanmalıdır. Ön işleme basamağı; filtreleme, bölütleme ve normalizasyon alt basamaklarından oluşmaktadır.



Şekil 2. Genel akış diyagramı ve ön işleme akış diyagramı

A. Filtreleme

Literatürde ivme verileri için en sık kullanılan filtre yapısı yürüyen ortalama filtredir. Adından da anlaşılacağı gibi bu filtrenin çalışma prensibi giriş sinyalinde belirlenen sayıdaki noktanın ortalamasının alınarak, çıkışa aktarılması şeklindedir. Denklem olarak gösterimi şu şekildedir:

$$y[i] = \frac{1}{M} \sum_{j=0}^{M-1} x[i-j] \quad (1)$$

Denklemdeki $x[]$ giriş sinyali, $y[]$ çıkış sinyali, M ise ortalaması alınan nokta sayısıdır.

Örneğin; 5 nokta için bu filtreyi uygulayıp 84 noktasının çıkış değerini bulmak için şöyle bir denklem kullanılmalıdır:

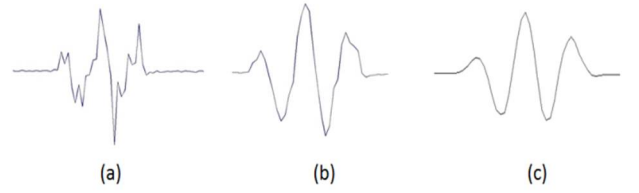
$$y[84] = \frac{x[84] + x[83] + x[82] + x[81] + x[80]}{5} \quad (2)$$

(2)'ye alternatif olarak şu şekilde bir denklem de kullanılabilir:

$$y[84] = \frac{x[82] + x[83] + x[84] + x[85] + x[86]}{5} \quad (3)$$

(3)'te noktalar çıkış noktasına göre simetrik olacak şekilde seçilmiştir [4].

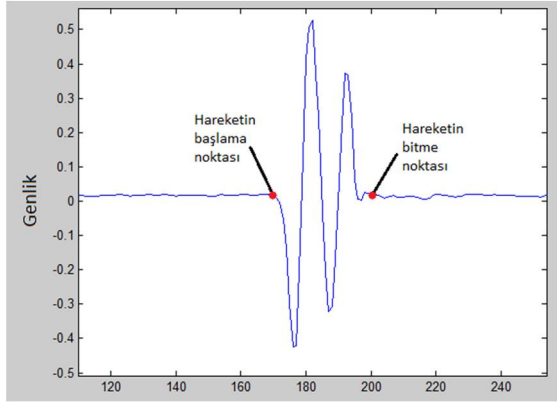
Şekil 3.(a)'da 'a' harfi için ivmeölçerden elde edilen X eksenli ham ivme sinyali gösterilmektedir. Sinyalin gürültülü yapısından dolayı uygulanan filtreleme (3)'teki gibi simetrik olacak şekilde) sonucunda (b)'deki sinyal elde edilmiştir. Filtrelenmiş sinyalde de giderilemeyen gürültüler bulunduğu için diğer çalışmalardan farklı olarak bu çalışmada aynı özellikli filtre (b)'de elde edilen sinyale 2. kez uygulanmış ve (c)'deki sinyal elde edilmiştir. Tanıma işlemi sonucunda 2 kez filtrelenen sinyallerden, daha yüksek oranlı bir doğru tanıma oranına ulaşıldığı görülmüştür.



Şekil 3. (a) Filtrelenmemiş, ham ivme verisi ('a' harfi) (b) Bir kez yürüyen ortalama filtre uygulanmış ivme verisi (c) İki kez yürüyen ortalama filtre uygulanmış ivme verisi

B. Bölütleme

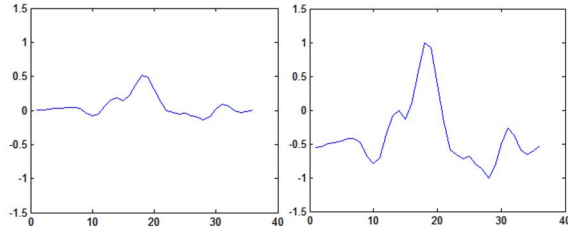
Yaptığımız çalışmanın veri toplama aşamasında her bir harf için 20 tekrar yapılmıştır ve tekrarlar arasında hareketsiz yaklaşık 3 saniye beklenmiştir. Sinyaldeki bu hareketsiz alanların kaldırılması için sinyalin genlik değerlerinden yararlanılır. Belirlenen eşik genlik değerleriyle hareketin olası başlama ve bitme noktaları belirlenir. Ancak bu belirlenen noktalar bazı durumlarda gerçekten kullanılması gereken noktalar olmayıp yazar kişinin elindeki ufak bir titremeden ya da sensörün oluşturduğu gürültüden dolayı oluşan bir sinyal de olabilir. Doğru sinyali bulmak ve hareketsiz alanları yok etmek için seçilmesi gereken sinyalin belirli bir süre devamlılığı olması gerektiği unutulmamalıdır. Hareketsiz alanlarda sinyalin genlik değeri 0 ya da 0'a çok yakın olduğundan örnek bir uygulamada 'b' harfi için hareketin yaklaşık olarak başlangıç ve bitiş noktaları Şekil 4'te görülmektedir.



Şekil 4. 'b' harfinin Y eksenindeki ivme sinyali

C. Normalizasyon

Bir önceki 'Bölütleme' kısmında da anlatıldığı gibi her bir harf için 20 tekrar yapılmıştır. Bu tekrarlar ideal bir sistemde genlik bakımından birbirine çok yakın sonuçlar vermelidir. Ancak pratikte bu şekilde ideal bir sistem üretmenin imkânsızlığından dolayı elde edilen tekrarlı sinyallerin ideale yaklaştırılması gereklidir. Bu sebeple her bir sinyalin maksimum ve minimum genlik değerlerinin sırasıyla +1 ve -1'e çekilerek normalize edilmesi gerekmektedir.



Şekil 5. Sırasıyla normalize edilmemiş ve normalize edilmiş 'a' harfi ivme verileri

III. DİNAMİK ZAMAN BÜKME METODU

Zaman serilerinde, Dinamik Zaman Bükme (DZB) metodu zaman ya da hız bakımından farklılık gösteren 2 sinyalin benzerlik ölçümü için kullanılan bir metottur. Örneğin; farklı kişilerin aynı harfi yazma hızları farklı olabilir ya da yazım esnasında farklı hızlanmalar ya da yavaşlamalar olabilir. İvme sensöründen elde edilen yazım hareketi verilerinin benzerlik değerleri de bu yöntemle hesaplanabilir.

Örneğin; R ve P sinyallerinin benzerliği bulunmak istenirse, sırasıyla uzunlukları n ve m olacak şekilde:

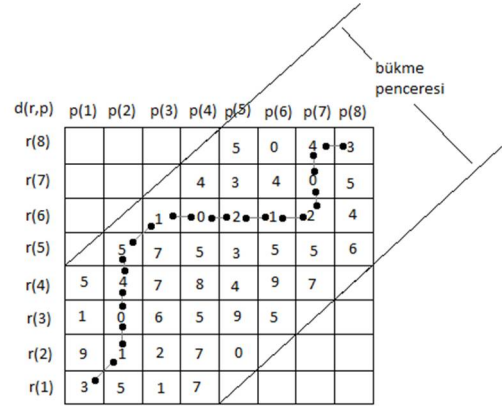
$$R = r_1, r_2, \dots, r_n$$

$$P = p_1, p_2, \dots, p_m$$

DZB metodunu kullanarak bu iki sinyali karşılaştırmak için öncelikle $n \times m$ 'lik bir matris oluşturulur. Bu matrisin i . ve j . elemanları r_i ve p_j noktaları arasındaki genlik bakımından uzaklık olan $d(r_i, p_j)$ 'yi verir. Kullanılan uzaklık ölçü denklemi tipik Öklid uzaklığıdır [5]:

$$d(r_i, p_j) = \sqrt{|r_i^2 - p_j^2|} \quad (4)$$

Matrisin her bir elemanı r_i ve p_j noktaları arasındaki karşılaştırmayı belirtir.



Şekil 6. Bükme yolu örneği

Şekil 6'daki ızgara yapı Öklid uzunluğuyla hesaplanan uzunluk matrisini ifade etmektedir. Izgara yapının her karesinin içinde benzerlik karşılaştırması yapılmak istenen iki sinyalin o kareye ait olan noktalarının uzaklık değerleri bulunmaktadır.

Şekil 6'da diyagonal olarak gösterilen bükme yolu (warping path) W ; iki sinyal arasındaki en optimal benzerliği gösterecek şekilde seçilmelidir. W 'nin k . elemanı $w_k = (i, j)_k$ şeklinde tanımlanmıştır ve şu şekilde gösterilir:

$$W = w_1, w_2, \dots, w_k, \dots, w_K$$

$$\max(m, n) \leq K < m + n - 1$$

Çok sayıda farklı bükme yolu bulunabileceğinden W 'nin bulunmasında çeşitli kısıtlamalar yapılmalıdır.

Bükme Penceresi: Bükme yolu adımlarını bulmaya çalışırken diyagonal çizgiden uzaklaşmamak için adımları sınırlandıran bükme penceresi kullanılır. Böylece karşılaştırması yapılan noktanın diğer sinyalde eşleşmemesi gereken çok farklı noktalara gitmesine engel olunabilir (Şekil 6).

Sınır Koşulları: Seçilmesi gereken bükme yolunun; uzaklık matrisinde diyagonal bir şekilde karşıt köşelerde başlayıp bitmesi gerektiğinden $w_1 = (1,1)$ ve $w_K = (m, n)$ baştan tanımlanmalıdır.

Süreklilik: Bükme yolundaki olası adımların komşu hücrelerden ilerlemesini sağlamak için her adımda yalnızca bir hücre ilerlemesi yapılmalıdır. $w_k = (a, b)$ olarak verildiğinde, $a - a' \leq 1$ ve $b - b' \leq 1$ koşulları için $w_{k-1} = (a', b')$ ile süreklilik sağlanabilir.

Monotonluk: W 'deki noktaların zamanda monotonik bir şekilde ilerlemesi için yani her adımda farklı bir hücreye ilerlemesini sağlamak için yine $w_k = (a, b)$

olarak verildiğinde, $a - a' \geq 0$ ve $b - b' \geq 0$ koşulları için $w_{k-1} = (a', b')$ olmalıdır.

Yukarıda anlatılan koşulları sağlayan çok sayıda bükme yolu vardır. Ancak ulaşılmak istenen yol sadece bükme maliyetini minimuma indiren yol olmalıdır.

$$DTW(R, P) = \min \left\{ \sqrt{\sum_{k=1}^K w_k / K} \right\} \quad (5)$$

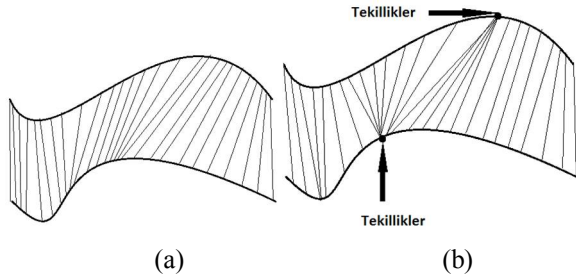
Paydadaki K bükme yollarının farklı uzunlukta olma olasılığına karşı kullanılmaktadır.[5]

Minimum maliyetli bükme yolu dinamik programlama ile etkili bir şekilde bulunabilir. Bunun için şöyle bir denklem kullanılır:

$$\gamma(i, j) = d(r_i, p_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \quad (6)$$

(6)'daki denklemde; $\gamma(i, j)$ komşu elemanların minimum uzunluğa sahip olanlarıyla oluşturulan ve mevcut hücrenin uzunluk değeri $d(i, j)$ ile bulunan birikimli uzaklık değeridir [5].

DZB yöntemi sinyallerin genlik değerlerini kullanıp benzerlik karşılaştırması yaptığı için bazı istenmeyen benzerlikler ortaya çıkabilmektedir. Bu istenmeyen benzerlik eşlemelerine 'tekillikler' adı verilmektedir. Adından da anlaşılacağı gibi sinyaldeki bir noktanın genlik değeri bakımından birbirine yakın olan, diğer sinyaldeki eşleşmemesi gereken başka noktalarla eşlenmesi ya da birden fazla noktanın tek bir noktaya eşleşmesi şeklindedir.



Şekil 7. Eşleşmesi gereken noktalar ve tekilliklere örnek

Şekil 7.(a)'da sinyallerde eşleşmesi gereken noktalar gösterilmiş, (b)'de ise DZB uygulamasından sonra eşleştirilen hatalı noktalar yani tekillikler gösterilmiştir.

Eamonn J. Keogh ve Michael J. Pazzani[5] yapmış oldukları çalışmada tekillikler sorununu çözmek için etkili bir yöntem olan Türevli Dinamik Zaman Bükmesi (TDZB) – Derivative Dynamic Time Warping (DDTW) yöntemini geliştirmişlerdir. Temel çalışma prensibi olarak DZB ile aynıdır ancak kullanılan veriler birbirinden farklıdır. DZB'de filtrelenmiş veriler doğrudan kullanılmaktadır ancak TDZB'de filtrelenen verilerin,

aynı DZB algoritmasını uygulamadan önce türevleri alınmaktadır.

$$D[r] = \frac{(r_i - r_{i-1}) + ((r_{i+1} - r_{i-1})/2)}{2} \quad (7)$$

Bu denklemde belirlenen nokta ve bir önceki noktadan elde edilen fark değeri ile belirlenen aynı noktanın bir öncesi ve bir sonrasındaki noktalardan elde edilen fark değerinin yarısının ortalaması kullanılarak türev hesaplanmaktadır[5].

IV. DENEYSEL ÇALIŞMALAR VE SONUÇLAR

3 eksenli ivme sensöründen elde edilen X, Y ve Z eksen ivme verilerinden sadece X ve Y eksen verileri tanıma amacıyla kullanılmıştır. Bunun nedeni yazma işlemin zemine paralel bir yüzey üzerinde yapılmasıdır. Z ekseninde yazma ile ilgili bir hareket yapılmadığından işlem yükünü artırmamak adına bu veriler ihmal edilmiştir.

Seçilen 13 harf (a, b, c, d, e, f, g, h, i, k, l, m, n) için, her veri kaydında 20'şer tekrarlı olarak alınan ivme verileri filtreleme, bölütleme ve normalizasyon işlemlerinden sonra tanıma algoritması yardımıyla karşılaştırılmıştır.

Seçilen bir karakter sinyali diğer 259 sinyal ile DZB algoritmasıyla karşılaştırılmış ve her bir karşılaştırmada seçilen sinyalin diğer sinyallerle elde edilen bükme yolu maliyeti kaydedilmiştir (leave-one-out). Yani her bir sinyal için 259 tane bükme yolu maliyeti kaydedilmiştir. Bu bükme yolu maliyetlerinden en küçük değerli olanın ait olduğu sinyal seçilen sinyalin eşleşmesi gereken sinyal olarak belirlenmiş ve bu işlem tüm sinyaller için uygulanmıştır.

DZB ile yapılan işlemlerde 260 karakter sinyalinden 5 tanesi hatalı olarak bulunmuştur. TDZB ile daha az sayıda hatalı tanıma oranı beklenirken bu çalışmada kullanılan sinyaller ile doğru tanıma oranı DZB ile benzer şekilde %98,08 olarak hesaplanmıştır.

KAYNAKÇA

- [1] Cho, S. J. , Oh, J. K. , Bang, W. C. , "Magic Wand: A Hand-Drawn Gesture Input Device in 3-D Space with Inertial Sensors", *Frontiers in Handwriting Recognition*, P 106-111, IEEE, 2004
- [2] Wang, J. S., Hsu, Y. L. and Liu, J. N. , "An Inertial-Measurement-Unit-Based Pen With a Trajectory Reconstruction Algorithm and Its Applications", *IEEE Transactions on Industrial Electronics*, 57(10):3508-3521, 2010
- [3] Zhang, S., Yuan, C., Zhang, Y. , "Handwritten Character Recognition Using Orientation Quantization Based on 3D Accelerometer", *Article No. 54, MobiQuitous*, 2008
- [4] Smith, S. W. , "The Scientist and Engineer's Guide to Digital Signal Processing", Second Edition, 1999
- [5] Keogh, J. E. and Pazzani, M. J. , "Derivative Dynamic Time Warping", *SIAM Conference on Data Mining*, Chicago, 2001