

Deneysel Mod Ayırıştırması Uygulanmış Yazma Hareket Bilgisi Kullanılarak El Yazısı Karakter Tanıma

Handwritten Character Recognition Using Empirical Mode Decomposition Applied Writing Movements

Esra TUNÇER ve Bilal Orkan OLCAY
Elektrik Elektronik Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü
İzmir, Türkiye
etuncer@iyte.edu.tr, bilalolcay@iyte.edu.tr

Mehmet Zübeyir ÜNLÜ
Elektrik Elektronik Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü
İzmir, Türkiye
zubeyirunlu@iyte.edu.tr

Özetçe—Bu çalışmada harflerin yazılma hareketlerinden yararlanılarak el yazısı karakter tanıma işlemi yapılmıştır. Yazma hareket bilgisini elde etmek adına 3 yönlü ivmeölçer kullanılmıştır. Mevcut çoğu sensörde olduğu gibi, ivmeölçerlerde istenen ivme bilgisinin yanında istenmeyen gürültü bilgisi de elde edilmektedir. Bu nedenle tanıma algoritmasından önce sinyaller ön işlemeden geçirilmeli ve gürültüden arındırılıp tanıma için uygun hale getirilmelidir. Bu amaca uygun olarak sinyallere Deneysel Mod Ayırıştırması (EMD) ve normalizasyon ön işleme adımları uygulanmıştır. Daha sonra sinyaller Dinamik Zaman Bükme (DTW) metodu ile karşılaştırılmış ve %91.92'lik doğru sınıflandırma oranı elde edilmiştir.

Anahtar Kelimeler — 3-yönlü ivmeölçer; deneysel mod ayırıştırma; dinamik zaman bükme; el yazısı karakter tanıma.

Abstract— In this paper, handwritten character recognition by using characters' writing movements is investigated. To obtain the information about writing movements a 3-axis accelerometer is used. Just like most of other sensors, 3-axis accelerometers give the actual movement signal as well as noise. Before the recognition step, all of the signals need to be preprocessed and the noisy parts need to be removed. So, Empirical Mode Decomposition (EMD) and normalization preprocessing steps are applied to the signals. Finally, the signals in the dataset are compared with Dynamic Time Warping for classification and accurate classification rate of 91.92% is obtained.

Keywords — 3-axis accelerometer; empirical mode decomposition; dynamic time warping; handwritten character recognition.

I. GİRİŞ

Karakter ve yazı tanıma, araştırmacıların özellikle son yıllarda sıklıkla çalıştığı konulardandır. Literatürde karakter tanıma denildiğinde daha çok karakter görüntüleri kullanılarak yapılan çalışmalara ulaşılmaktadır. Bu tarz çalışmalarda uzun yıllar bilgisayar tanımlı karakterler üzerine çalışılmıştır. Son

yıllarda el yazısı ile yazılmış karakterler üzerine yapılan çalışmalar da hız kazanmıştır ancak bu çalışmaların birçoğunda yazma anındaki hareket bilgisi bulunmadığından kişiden kişiye farklılık gösteren karakterlerde bazı tanıma sorunları ortaya çıkmaktadır. Aynı zamanda şeklen aynı tarz kullanılsa bile hiçbir kullanıcı bir karakteri birebir aynı şekilde yazamaz. Bu da kullanıcı tanımlı olmayan çalışmalarda ortaya çıkan en büyük problemdir. Karakterlerin yazılma hareketlerinin elde edilebilmesi adına kamerayla yazma hareketlerinin takibi, literatürde görüntü işleme çalışmaları için kullanılan metodların temeli oluşturmaktadır. Ancak yaptığımız bu çalışmada karakterlerin görüntü bilgilerinden değil, karakterlerin ya da harflerin yalnızca yazılma hareketlerinden yararlanılarak sınıflandırma işlemi yapılmıştır. Görüntü işleme temelli çalışmalarda kullanılan, kamerayla hareket takibi yerine literatürde de sıklıkla karşılaşılan 3 eksenli ivmeölçer ile hareket bilgisi ölçümü bu çalışma için tercih edilmiştir.

İvmeölçerler karakter hareket bilgisinin elde edilmesinin yanında vücut hareket bilgilerinin çeşitli uygulamalar için toplanması amacıyla da sıklıkla kullanılmaktadır. Bu bağlamda karakter yazma hareketleri de vücut hareketleri gibi düşünülebilir. Ancak bu hareketler daha karmaşık ve sınıf sayısı bakımından çoğu çalışmaya göre daha çeşitlidir.

Sung-Jung Cho yapmış olduğu çalışmada 13 adet farklı sınıfa ait, basit vücut hareketlerini ivme ve açısal hız bilgilerinden yararlanarak sınıflandırmıştır [1]. Sung-Do Choi ve Shiqi Zhang, sırasıyla, 10 adet rakamın ivme bilgilerini [2] ve ivme bilgisinin hız ve konum bilgilerine dönüşümünü [3] kullanmışlardır. Klingmann ise çalışmasında iPhone(Apple Inc.)'un içinde bulunan ivmeölçer ile 5 farklı hareketin (halka, kare, üçgen, z şekli ve bowling topu atma hareketi) ivme değerlerini elde edip sınıflandırmasını yapmıştır [4]. Bazı çalışmalarda ise hareket ivme bilgileri kullanıcıya söylenen kısıtlayıcı başlangıç şartları ile elde edilmiştir. Örneğin; Meenaakumari'nin çalışmasında kullanıcılara karakter ve

rakamları dijital şekilde yazmaları söylenmiştir. Böylece Şekil.1’de gösterildiği gibi 4 farklı yönün (aşağı, yukarı, sağ ve sol) kombinasyonları elde edilmiş ve tanıma işlemleri aslında sadece bu 4 farklı sınıf üzerinden yapılmıştır [5].

1	2	3	4
←	→	↑	↓

Şekil. 1. Meenaakumari’nin çalışmasındaki sınıflar [5]

Yapmış olduğumuz bu çalışmada, İngiliz alfabesindeki 26 harfin yazma anındaki hareket bilgileri 3 eksenli ivmeölçer ile elde edilmiştir. Bu çalışmanın amaçlarından biri de en az sayıda mevcut bilgiyle en yüksek doğrulukta tanıma olduğundan ve normal yazma işlemi bir düzlem üzerinde yapıldığından Shashidhar Patil’in çalışmasındaki [6] aksine yazma işlemi 3 boyutta değil bir düzlem üzerinde ve ivmeölçerden gelen ivme bilgilerinin sadece 2 eksen kullanılarak yapılmıştır. 3’üncü eksenin ileride yapılacak çalışmalarda kelime içerisindeki harfleri zamansal düzlemde ayırmada kullanılması planlanmaktadır.

Piyasadaki çoğu sensör gibi ivmeölçerler de arzu edilen hareket bilgisinin yanında sensörün iç gürültüsünü de vermektedir. Bunun yanında elde edilen sinyaller hareket esnasında oluşan titreşim ve yazma yüzeyinin bozukluğundan kaynaklı bir gürültüye de sahiptir. Bu gürültüler elde edilmek istenen hareket bilgisini bozduğundan tanıma algoritmalarında yanlış sınıflandırmalara sebep olabilmektedir. Bu çalışmada sinyallerin bu gürültülerden arındırılması adına Deneysel Mod Ayırıştırması (EMD) kullanılmıştır. Bu metot ikinci bölümde detaylı olarak anlatılacaktır.

Kullanıcılar aynı harfi yazıyor bile olsa yazma hızı ve ivmelenmesi bakımından aralarında farklılıklar olacaktır. Bu da sinyaller arasında genlik farklılıklarına sebep olmaktadır. Bu farklılıkları ortadan kaldırmak ve sinyallerin benzerliklerini artırmak için çalışmada normalizasyon işlemi uygulanmıştır. Yine aynı şekilde kullanıcılar aynı harfi yazıyor bile olsa bir harfin tamamlanması için geçecek olan süre kişiden kişiye değişmektedir. Bu durum, aslında tek bir kullanıcı için de birebir aynı süreler içinde yazma işlemini tamamlayamayacağından, geçerlidir. Bu sebeple sinyallerin boyları arasında da büyük farklılıklar olacaktır. Tanıma algoritması bu şartlara uygun olacak şekilde seçilmelidir. Sinyaller arasındaki benzerlik hesaplamasını, seçilen 2 sinyalin tüm noktaları arasında uzaklık hesaplaması ile yapan Dinamik Zaman Bükme (DTW) metodu bu çalışmanın tanıma tarafı için kullanılmıştır.

II. ÖNİŞLEME

A. Deneysel Mod Ayırıştırması (EMD)

Hilbert-Huang dönüşümü Huang tarafından 1998 yılında önerilen, sinyal ve verileri İç Mod Fonksiyonları (Intrinsic Mode Functions-IMF) adı verilen parçalara ayıran bir metottur.

IMF’lerin bulunması ve çıkarılması işlemine eleme işlemi denilmektedir. Bu işlemde ilk önce sinyalin tüm yerel maksimum-minimum noktaları belirlenir ve iki adet yukarı zarf ve aşağı zarf adı verilen fonksiyon bu noktalar kullanılarak ve ara değerlendirme (interpolation) ile oluşturulur. İkinci olarak, bu iki zarf fonksiyonunun ortalaması alınarak orijinal sinyalden daha düşük frekansa sahip bir bileşen elde edilir. Zarf ortalamalarının orijinal sinyalden çıkarılması ile de en yüksek frekanslı bileşen (IMF1) elde edilir. Elde edilen dalganın maksimum ve minimum nokta sayılarının ve sıfırdan geçiş sayılarının arasında yalnızca 1 fark varsa ve dalganın yerel

ortalaması sıfır ise bu dalga IMF olarak tanımlanmaktadır. Eğer bu şartlar ilk tekrarda sağlanamazsa aynı işlemler şartlar sağlanana kadar tekrarlanmalıdır [6].

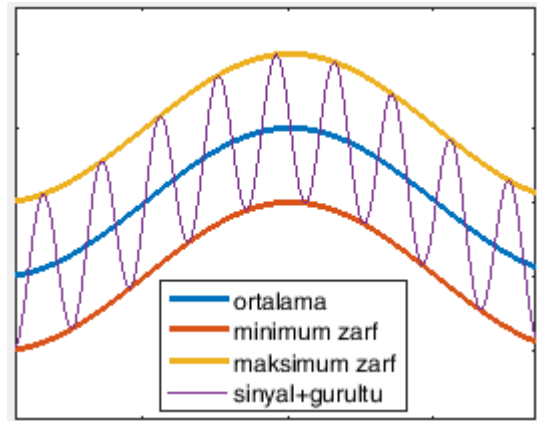
Şekil.2’deki sarı çizgi yerel maksimum noktalarının oluşturduğu yukarı zarfı, kırmızı çizgi ise yerel minimum noktalarının oluşturduğu aşağı zarfı temsil etmektedir. Mavi ile gösterilen çizgi yukarı ve aşağı zarfların ortalamasıyla oluşturulan dalgayı temsil etmektedir. Aynı zamanda bu mavi dalga gürültülü sinyalin alçak geçiren filtre ile filtrelenmiş hali gibi düşünülebilir.

B. Normalizasyon

Kullanıcı tanımlı olmayan çalışmalarda kullanıcıların yazma biçimleri arasında farklılıklar olması kaçınılmazdır. Aynı karakter yazılsa bile yazma hızı ve ivmelenmesi arasında farklılıklar olacaktır. Bu farklılıklar sinyallerin şekilsel bakımdan yapılarını büyük ölçüde etkilemese bile ivmelenme değerleri ve dolayısıyla sinyallerin genlik değerleri arasında tanıma adımı açısından olumsuz yönde etkileyecek farklılıklara sebep olacaktır. Sonuç olarak eşleşmesi gereken sinyaller yerine hatalı eşleşmeler oluşacaktır. Bu sebeple sinyallerin genlik değerleri eşitlik 1.’de verilen denklemle belirli bir aralığa çekilmelidir. Bu çalışmada maksimum değer olarak +1 ve minimum değer olarak -1 genlik değerleri seçilmiştir.

$$y = m + (x - A) * (n - m) / (B - A) \quad (1)$$

Burada x normalize edilecek olan sinyal, m minimum limit, n maksimum limit, A sinyalin maksimum genlik değeri ve B sinyalin minimum genlik değeridir.



Şekil. 2. Temsili aşağı-yukarı zarflar ve ortalamaları

C. Hareket aralığının bekleme aralığından ayrılması

Kullanıcı sinyal kaydı yapılırken yazma işlemi yapılan kadar belirli bir süre boyunca zemin üzerinde sabit bir şekilde beklemektedir. Kaydın sonunda elimizde hem hareketin olduğu hem de beklemenin olduğu kısımlar bulunmaktadır. Her sinyalde bulunacak olan ve benzerliği yüksek olan bekleme aralıkları hatalı tanımlara sebebiyet vermektedir. Bu sebeple sinyallerin sadece hareketin olduğu kısımları kullanılmalıdır. Burada kullanabileceğimiz tek şey sinyallerin genlik değerleridir. Ancak bekleme anında genlik değeri çok küçük olmalı diyerek küçük genlikli parçaları sinyalden silerek o zaman gerçek sinyalde de bazı önemli parçaların kaybedildiği gözlemlenmiştir. Bu sebeple sinyalin genlik değerleri değil sinyal örnekleri arasındaki genlik farklılıklarına bakmamız daha uygun olur. Eğer genlik farkı seçilen bir eşik değerinin üzerindeyse bu noktalar sinyalin olası başlangıç ve bitiş değerleri olarak seçilmelidir (Şekil.3). Son olarak bulunan sinyalin uzunluğunun belli bir örnek değerinden fazla olup olmadığına bakılmalıdır.

III. DİNAMİK ZAMAN BÜKME ALGORİTMASI

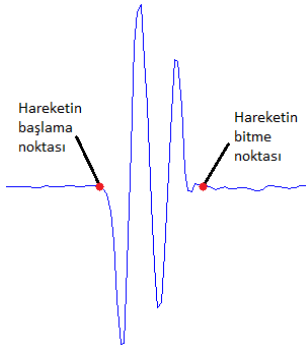
Dinamik Zaman Bükme (DTW); zamansal eksen ya da faz farklılığına sahip sinyallerin arasındaki benzerlik miktarlarının bulunması için literatürde sıklıkla kullanılan bir yöntemdir. Bu çalışma için düşünce olursak, kullanıcıların aynı harfi yazma hızlarında ve ivmelenme değerlerinde farklılıklar olması kaçınılmazdır. Bu sebeple aslında hareketin aynı yerine ait olan sinyal parçacıkları farklı zaman değerlerine kaymış şekilde elde edilmektedir. İvme ölçer ile elde edilen yazma hareketleri arasındaki benzerlik miktarı için bu yöntem kullanılmaktadır [7] [8].

Uzunlukları n ve m olan, U ve V karakterlerinin ivmelenme sinyalleri

$$U = u_1, u_2, \dots, u_n$$

$$V = v_1, v_2, \dots, v_m$$

şeklinde tanımlanırsa, ilk olarak $n \times m$ boyutunda uzaklık matrisi oluşturulur. Uzaklık matrisinin i . satır ve j . sütunundaki değer, sinyallerin u_i ve v_j noktaları arasındaki uzaklık değeridir. Uzaklık matrisi hesaplamasında Denklem (2)'deki klasik Öklid uzaklığı tek boyutlu sinyaller için kullanılmaktadır. Ancak bu çalışmada kullandığımız uzaklık ölçü denklemi iki boyutlu sinyallere uygulanan Denklem (3)'teki Öklid uzaklığıdır [8]:



Şekil. 3. Hareketin başlama ve bitme noktaları [7]

$$d(u_i, v_j) = \sqrt{|u_i^2 - v_j^2|} \quad (2)$$

$$d(u_i, v_j) = \sqrt{(u_{x_i} - v_{x_j})^2 + (u_{y_i} - v_{y_j})^2} \quad (3)$$

Şekil.4 Öklid uzaklığıyla hesaplanan temsili bir uzaklık matrisidir. Köşegensel şekilde ve siyah renk ile gösterilen yol (bükme yolu - w), sinyaller arasındaki ideale en yakın eşleşme noktalarını sağlayacak ve toplam uzaklık değeri minimum olacak şekilde seçilmelidir. Ancak benzer şartları sağlayan birden fazla w bulunabileceğinden bazı kısıtlamaların uygulanması kaçınılmazdır [9]:

- Bükme yolu köşegensel çizgi etrafından seçilmelidir. Bunu sağlamak adına bükme penceresi kullanılmalıdır. Pencere kullanımı sadece genlik değerlerinin benzerliği fazla diye hatalı 2 noktanın eşleştirilmesini engeller (Şekil.4'teki kırmızı yol gibi). Aynı zamanda bükme penceresi, uzaklık matrisinin sadece bu pencere içinde kalan noktalarında uzaklık hesaplaması yapma yeterliliğini sağladığından işlem hızını artırıcı bir etkisi de bulunmaktadır [7].
- Bükme yolunun başı ve sonu, uzaklık matrisinin köşegensel karşıt noktalarında olmalıdır [7].
- Bükme yolunun zamanda monotonik bir şekilde ilerlemesi ve her yeni eşleşmenin bir önceki eşleşmenin komşu hücreleriyle yapılması sağlanmalıdır [7].

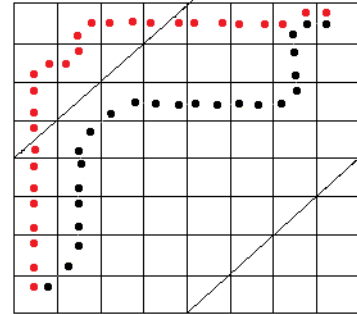
Kısıtlamaların ardından bile aynı şartları sağlayan birçok bükme yolu seçeneği ile karşılaşılabilir. Seçilmesi gereken yol, yol üzerindeki uzaklık değerleri toplamı en küçük olandır. Denklem (4) seçim işlemi için kullanılabilir.

$$DTW(U, V) = \min \left\{ \sum_{k=1}^K w_k / K \right\} \quad (4)$$

K , karşılaştırılan sinyal çiftlerinden elde edilen bükme yollarının farklı uzunlukta olması ihtimaline karşı kullanılmalıdır. Uzaklık değer toplamı en küçük olan yol

$$s(i, j) = d(u_i, v_j) + \min\{s(i-1, j-1), s(i-1, j), s(i, j-1)\} \quad (5)$$

dinamik programlamasıyla bulunabilir. $s(i, j)$; minimum değerli komşu ve $d(u_i, v_j)$ 'nin toplamı ile oluşturulan toplam (birikimli) uzaklık değerini verir [9].



Şekil. 4. Bükme yolu örneği

IV. DENEYSEL ÇALIŞMALAR VE SONUÇLAR

Bu çalışmada İngiliz alfabesindeki 26 adet küçük harfi tanıma işlemi yapılmıştır. Veri setinde her bir karakter için 20 adet sinyal örneği bulunmaktadır. Tanıma için karakterlerin yazılma hareketlerinden yararlanılmış ve bu hareket bilgileri 3 eksenli bir ivmeölçer ile elde edilmiştir. Normal yazma işlemi bir düzlem üzerinde yapıldığından burada da yazma düzlemine karşılık gelen x ve y eksenleri kullanılmıştır. Toplamda elimizde 1040 adet sinyal bulunmaktadır. Üçüncü eksen olan z ekseninin ise ileride yapacağımız çalışmalarda bir kelime ya da cümle yazımında her bir harf arasındaki boşluğu belirlemede kullanılması planlanmaktadır. Literatürdeki bazı çalışmaların aksine, iki eksen kullanılmasının bir diğer sebebi ise yazma hareketi ile ilgili minimum seviyede bilgiyle nasıl bir sonuç elde edileceğinin gözlemlenmek istenmesidir.

Piyasadaki birçok sensörde olduğu gibi kullandığımız ivmeölçer de hareketin ivme bilgisinin yanı sıra sensördeki iç gürültüyü ve hareket anındaki titreşimleri de vermektedir. Bu gürültüler ile neredeyse hiçbir tanıma metodu tatmin edici sonuç veremeyeceği için öncelikle veri setindeki tüm sinyallerin düzenlenmesi ve gürültüden yeterli seviyede arındırılması gerekmektedir. Bu çalışmada filtreleme kısmı için literatürde daha önce bu tarz bir çalışma için kullanıldığını görmediğimiz ve bir önceki çalışmamızda [7] da kullanmadığımız Deneysel Mod Ayırıştırma metodu kullanılmıştır. Bu metot sinyalin yerel maksimum ve minimum noktalarını belirleyip bunlar üzerinden sırasıyla yukarı ve aşağı zarf adı verilen iki yeni fonksiyon oluşturmaktadır. Daha sonra bu fonksiyonların ortalaması alınarak düşük frekanslı bir sinyal ve gürültülü sinyalden de ortalama sinyalini çıkararak yüksek frekanslı bir sinyal (IMF1) elde edilmektedir. Bu işlem yeterli seviyede filtreleme sağlanıncaya kadar ya da kullanma amacına göre istenen sinyal bileşeni elde edilene kadar tekrar edilebilir. Biz bu çalışmanın ilk denemelerinde iki tekrarın (IMF2) frekans bölgesi bakımından sinyali istediğimiz seviyede filtrelediğini görmüştük. Ancak zaman bölgesindeki filtreleme sonuçlarına baktığımızda birçok sinyalin şekilsel bakımdan bozulduğunu gördük. Bu sebeple sinyallere EMD'yi bir kez (IMF1) uyguladık ve elde edilen yüksek frekanslı bileşeni gürültülü sinyalden çıkardığımızda şekilsel bozulmaların giderildiğini gördük.

Kullanıcıdan kullanıcıya yazma hızı ve ivmelenmeleri değişiklik gösterdiğinden elde edilen sinyaller genlik ve faz farkı bakımından farklılıklar göstermektedir. Genlik farklılıklarını gidermek adına sinyaller +1 ve -1 genlik değerleri arasına normalize edilmiştir. Normalizasyon işleminin ardından sinyalin hareket olmayan, bekleme kısımlarından ayrılması işlemi yapılmıştır.

Tüm ön işleme adımlarının ardından elimizde boyları ve ivmelenmesi farklı sinyaller bulunmaktadır. Bu sinyallerin birbirlerine olan benzerlik miktarlarını bulmak için Dinamik Zaman Bükmesi (DTW) metodu bu çalışmanın tanıma kısmı olarak kullanılmıştır. Bu metot, boyu farklı olan ve aralarında faz farkı olan sinyallerin birbirlerine olan benzerliklerini Öklid mesafe ölçümü temeline dayanan bir teknikte bulunmaktadır. Tek bir kullanıcıdan elde edilen sinyallerle oluşturduğumuz veri setinden seçilen bir sinyal bir tanesini dışarda bırak

(leave-one-out) metoduyla veri setinde geriye kalan tüm sinyallerle karşılaştırılmış ve karşılaştırılmak üzere seçilen sinyal minimum toplam uzaklık değerini sağlayan sinyal ile etiketlenilmiştir.

EMD ile iki tekrarlı (IMF2 kullanılarak) olarak filtrelenen sinyaller yüksek hatalı sınıflandırma miktarıyla %68.07 şeklinde doğru tanıma oranı vermiştir. Ancak EMD ile bir kez (IMF1 kullanılarak) filtrelenen sinyaller ile %91.92'lik bir doğru tanıma oranına ulaşılmıştır. Bu tanıma yüzdeleri bize DTW için sinyalin şekilsel yapısının korunması gerektiğini göstermiştir. Ayrıca tek tekrar (IMF1) ile filtrelenen sinyallerin doğal olarak gürültü miktarları 2 tekrarlardan (IMF2) fazla olduğu için; DTW metodunun buradaki gibi gürültüden tam anlamıyla arındırılmayan sinyaller üzerinde de son derece tatmin edici sonuçlar verdiği sonucuna da ulaşılabilir.

KAYNAKLAR

- [1] Sung-Jung Cho, J. K. O., Won-Chul Bang, Wook Chang, Eunseok Choi, Yang Jing, Joonkee Cho, Dong Yoon Kim. (2004). *Magic Wand: A Hand-Drawn Gesture Input Device in 3-D Space with Inertial Sensors*. Paper presented at the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004).
- [2] Shiqi Zhang, C. Y., Yan Zhang. (2008, 2008-07-21). *Handwritten character recognition using orientation quantization based on 3D accelerometer*. Paper presented at the Mobiquitous '08 Proceedings of the 5th Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services, Belgium.
- [3] Sung-Do Choi, A. S. L., Soo-Young Lee. (2006). *On-Line Handwritten Character Recognition with 3D Accelerometer*. Paper presented at the International Conference on Information Acquisition, China.
- [4] Klingmann, M. (2009). *Accelerometer-Based Gesture Recognition with the iPhone*. (MSc.), Goldsmiths University of London, London.
- [5] Meenaakumari, M. (2013). *Mems Accelerometer Based Hand Gesture Recognition*. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 2(5), 1886-1892.
- [6] Kim, D., & Oh, H. S. (2009). EMD: a package for empirical mode decomposition and Hilbert spectrum. *The R Journal*, 1(1), 40-46.
- [7] Tunçer, E., & Ünlü, M. Z. (2016, May). Accelerometer based handwritten character recognition using Dynamic Time Warping method. In *Signal Processing and Communication Application Conference (SIU), 2016 24th* (pp. 1805-1808). IEEE.
- [8] Tunçer, E. (2017). *Accelerometer based handwritten character recognition using dynamic time warping* (Master's thesis, İzmir Institute of Technology).
- [9] Eamonn J. Keogh, M. J. P. (2001). *Derivative Dynamic Time Warping*. Paper presented at the First SIAM International Conference on Data Mining