**PROJE ADI: KONUŞMADAN METNE DÖNÜŞÜM**

**GRUP ADI: MİCROMEGA**

**GRUP ÜYELERİ:**

**16220054 - KAYAHAN MELİH ALTINÇIBIK**

**16220014 - BURAK ELHAMAN**

**16220020 - NURULLAH ALAGÖZ**

**15220055 - YUNUS YILMAZ**

**16220062 - MEHMET DAŞÇİ**

**14210034 - AYETULLAH AKDOĞAN**

**12.12.2019**

Bilgisayarların gündelik yaşamın bir parçası olmasıyla birlikte insan-bilgisayar etkileşimi önem kazanmaya başlamıştır. Geleneksel insan-bilgisayar ara yüzleri olan klavye, fare gibi aygıtlar yaygın olarak kullanılmalarına rağmen kullanıcı ile bilgisayar arasındaki bilgi ve komut akışını kısıtlamaktadır. Ayrıca geleneksel ara yüzlerin büyük bir kısmı motor yetmezlikleri olan fiziksel engelliler tarafından kullanılamamaktadır. Bu yetersizliklere bir çözüm olarak bilgisayarın konuşma sinyali ile kontrol edilmesi önerilebilir. Ancak, lehçe ve dilsel farklılıklar gibi nedenlerden dolayı, geniş bir kelime haznesine sahip sistemlerin tasarlanması son derece zordur. Yine de kısıtlı sayıda komutla gerçekleştirilecek uygulamaların başarımları tatmin edici düzeylerde olabilmektedir.

Konuşma tanıma, konuşma sinyallerinden otomatik olarak dilsel bilgilerin çıkarılması olarak tanımlanır. Konuşma tanıma sistemlerinin çalışma ilkesi kabaca, giriş verisinin daha önce kaydedilmiş şablonlarla karşılaştırılmasına dayanır. Bu şablonlar fonem düzeyinde ya da kelime düzeyinde düzenlenmiş olabilir. Bu karşılaştırma sonucunda verinin en çok benzediği şablon o verinin simgesel gösterimi olarak kabul edilir. Ses sinyallerinin doğrudan karşılaştırılması son derece uzun işlemler gerektirdiğinden ses sinyalleri üzerinde bilgi miktarını azaltmadan veri boyutunu azaltacak bir ön işlem yapılması gerekmektedir. Bu ön işlem öznitelik çıkarımı olarak adlandırılır. Alınan ses verisinden öznitelik vektörleri çıkarılarak elde edilen öznitelikler daha önce hesaplanarak etiketlenmiş olan şablonlarla karşılaştırılır. Burada elde edilen öznitelik vektörlerinin ses verilerini en iyi şekilde temsil etmesi, işlenebilir oran ve büyüklükte olması ve ayırt edici özelliklere sahip olması gerekmektedir. Literatürde konuşma tanıma uygulamalarında kullanılabilecek birçok öznitelik tanımlanmıştır. Bu özniteliklerden Mel Frekans Kepstral Katsayıları(MFCC) kullanılacaktır.

Konuşma tanıma sistemleri, ayrık/sürekli, kişiye bağımlı/kişiden bağımsız, kelime tabanlı/fonem tabanlı gibi çeşitli kategorilerde sınıflandırılabilirler. Ayrık konuşma tanıma sistemlerinde, kısa aralıklarla seslendirilen sözcüklerin tanınması amaçlanırken sürekli konuşma tanıma sistemlerinde ise ara verilmeden seslendirilen sözcüklerin tanınması amaçlanır. Kişiye bağlı sistemlerde referans şablonları tek bir kişi için oluşturulurken, kişiden bağımsız sistemler tanıma işlemini konuşmacının kimliğinden bağımsız olarak gerçekleştirilir. Kelime tabanlı konuşma tanıma sistemleri, tanıma için öngörülen en küçük birim olarak kelimeyi kabul eder. Fonem tabanlı konuşma tanıma sistemleri ise tanıma için fonemlerin en küçük birim olarak kabul edildiği sistemlerdir. Bütün konuşma sesleri, farklı frekans değerlerine sahip sinüs dalgalarının doğrusal birleşiminden oluşur. İnsan sesinin frekans değerleri 300Hz-3300Hz arasında değişmektedir. Nyquist Teoremine göre ses frekansının iki katı ve daha büyük örnekleme frekansı ile etkin bir örnekleme yapılır.

Projemizde konuşulan kelimelerin Braille alfabesine dönüştürülmesi üzerinde durulmuştur. Braille alfabesi görme engellilerin okuyup yazması için 1821 yılında Louis Braille tarafından geliştirilmiştir. Projemizin yazıcı kontrolünde kullanılan yazılımı ve donanımın bir kısmı değiştirilerek normal yazılar için yani yazma engeli olan ve görebilen insanlar için de kullanılabilir. Projemizde ayrık konuşma tanıma kullanılarak kısa aralıklı seslerin sözcüklerinin tanınması, kelime tabanlı olarak ve kullanacağımız yapay zeka sayesinde kişiden bağımsız olarak kullanılması amaçlanmaktadır.

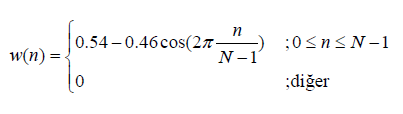
**İş Planı Çizelgesi**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Yapılması Planlanan İş** | 1. Hafta | 2. Hafta | 3. Hafta | 4. Hafta | 5. Hafta | 6. Hafta | 7. Hafta | 8. Hafta | 9. Hafta | 10. Hafta |
| Ses tanima - konuşmadan metne dönüşüm hakkında literatür taraması |  |  |  |  |  |  |  |  |  |  |
|
| Yöntemin belirlenmesi ve teorik araştırma |  |  |  |  |  |  |  |  |  |  |
| Mikrodenetleyecinin seçilmesi ve yazılımının hazırlanması |  |  |  |  |  |  |  |  |  |  |
| Devre tasarımı ve malzeme temini |  |  |  |  |  |  |  |  |  |  |
| Makinenin ve baskı devrenin prototipinin oluşturulması |  |  |  |  |  |  |  |  |  |  |
| Makine elamanlarının yazılımsal işlemlerinin yapılması |  |  |  |  |  |  |  |  |  |  |
| Simülasyonun yapılması ve prototipin test edilmesi |  |  |  |  |  |  |  |  |  |  |
| Prototipin nihai ürüne dönüştürülmesi |  |  |  |  |  |  |  |  |  |  |
| Satış ve pazarlama |  |  |  |  |  |  |  |  |  |  |

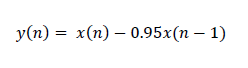
**İş Dağılımı Çizelgesi**

|  |  |  |
| --- | --- | --- |
|  | **Yapılacak İş** | **Kim Tarafından Yapılacağı** |
| 1 | Projede kullanılacak Mikrodenetleyicinin belirlenmesi ve çevre birimleri ile iletişiminin araştırılması | Burak Elhaman Nurullah Alagöz |
| 2 | Mikrodenetleyici yazılımının yazılması | Burak Elhaman Kayahan Melih Altınçıbık |
| 3 | Yazıcı bölümünün yapılması ve yazılımının yazılması | Nurullah Alagöz Kayahan Melih Altınçıbık  Ayetullah Akdoğan |
| 4 | Teorik kısımların araştırılması ve grup üyeleri ile paylaşılması | Yunus Yılmaz Mehmet Daşçi Burak Elhaman |
| 5 | Projede kullanılacak malzemelerin temininin ve piyasa araştırmasının yapılması | Nurullah Alagöz |
| 6 | Grup koordinasyonu, toplantı ve çalışma saatlerinin ayarlanması | Mehmet Daşçi |

Projemize başlarken mikrodenetleyici kullanmayı düşünmüştük. Bunun için mikrodenetleyicilerin mikrofon ve bluetooth ile etkileşimlerini araştırıp kullanmaya çalıştık. Ancak sonradan projemiz için hafıza ve ram ihtiyacının mikrodenetleyicilerin çoğunun karşılamayacağını farkettik. Bu nedenle Matlab kullanmaya karar verdik. Matlab kullanmamızın bir diğer faydası da içerisinde kullanabileceğimiz yapay zeka sayesinde projemize kişiden bağımsız olarak kullanma imkanı sağlamasıdır. Matlab’daki kodumuzda ilk olarak bilgisayarın ses kaydedicisinden oluşturduğumuz ses kaydını programa ekliyoruz. Sonra alçak geçiren FIR filtre ile filtreliyoruz. Buradaki frekans aralıklarını deneme amaçlı olarak değiştiriyoruz. Ancak teorik olarak 300 ile 3300Hz arasında olması gerekiyor. Bu işlemlerden sonra çerçeveleme işlemi geliyor. Her bir çerçeve boyutunun 20 ile 30 ms arasında olması gerekmektedir. Biz de 25ms olarak ayarladık ve çerçeve boyutumuzun 1200 örnek olduğunu gördük. Ancak bir çerçeve alınırken pencereleme işlemi için bir önceki ve bir sonraki çerçeveden 600 örnek alınır ve örtüşme sağlanır. Böylece sinyalin karakteristiklerini taşıyan katsayıların hesaplanması kolaylaşır ve sürekli bir sinyal elde edilir. Çerveleme işlemi de bittikten sonra pencereleme işlemi gelmektedir. Pencereleme işlemi için Hamming pencereleme yöntemini kullandık. Matematiksel olarak çerçeveleme işlemi özgün sinyali bir dikdörtgen pencere ile çarpmaya denktir. Hamming penceresinin matematiksel ifadesi aşağıdaki gibidir.



Bu işlemlerin ardından ön vurgu filtresi işlemini yaptık. Ön vurgu filtre işlemi normalde ilk kısımda yapılıyor ancak grafiklerin daha sağlıklı gözlemlenebilmesi açısından pencerelemeden sonra kullandık. Bu filtrenin amacı sinyalin yüksek frekans spectrumuna ilişkin enerjisinin arttırılmasıdır. Ön vurgu filtresinin zaman bölgesindeki transfer fonksiyonu aşağıdaki eşitlikte gösterilen FIR (Sonlu Dürtü Tepkisi) filtresi ile yapılır.



**Matlab Program Kodu**

close all; clear all;

%yeni deneme(fft fonksiyon ile)

[y,fs]= audioread('C:\Users\User\Desktop\proje\Ses kayıtları\ses.wav');

%length(a)/fs; kaç sn///sound(y,fs); çalma///sound(y,fs\*2); iki kat hızlı çalma

% 1. AŞAMA FİLTRELEME

FIR=fdesign.lowpass('Fp,Fst,Ap,Ast',300,10000,0.5,60,fs);%300Hz ile 10000Hz arasında olan frekansları alır

D=design(FIR);

a=filter(D,y);

preemph = [1 -0.95];%ön vurgu parametreleri

% 2. AŞAMA ÇERÇEVELEME

f\_d = 0.025; % çerçeve süresi

f\_size = f\_d \* fs; %çerçeve boyutu 1200

for c=1:(length(y)/f\_size)\*2;

for k=1:f\_size;

fr(k,c) = a(k+(c-1)\*(f\_size/2)); %Çerçeveleme işlemi 600 örnek önceden ve 600 örnek sonradan alır,

end

fr\_win(c,:) = fr(:,c) .\* hamming(length(fr(:,c)));%Hamming pencereleme işlemi

e(c,:)=filter(preemph,1,fr\_win(c,:));%ön vurgu uygulama

end

subplot(5,1,1)

plot(y)

xlabel('Zaman(sn)');ylabel('Genlik');title('Sesimizin Görüntüsü');

subplot(5,1,2)

plot(a)

xlabel('Zaman(sn)');ylabel('Genlik');title('Filtrelenmiş hali');

subplot(5,1,3)

plot(fr(:,135))

xlabel('Zaman(sn)');ylabel('Genlik');title('Bir Çerçeve');

subplot(5,1,4)

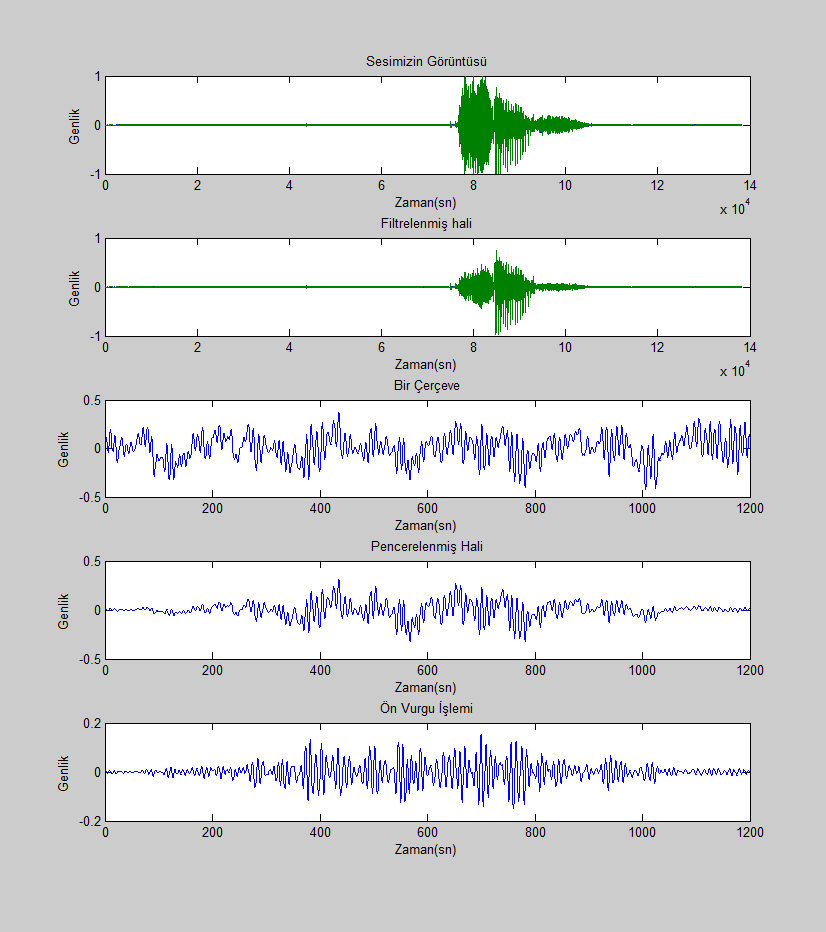
plot(fr\_win(135,:))

xlabel('Zaman(sn)');ylabel('Genlik');title('Pencerelenmiş Hali');

subplot(5,1,5)

plot(e(135,:))

xlabel('Zaman(sn)');ylabel('Genlik');title('Ön Vurgu İşlemi');

**Program Çıktısı**

**Malzeme Listesi**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Malzeme adı** | **Adet** | **Model** | **Adet fiyatı** | **Fiyat** |
| Mikrofon | 1 | Max4466 | ₺14,4 | ₺14,4 |
| Step motor | 7 | 28-BYJ 48 | ₺13,1 | ₺91,7 |
| Jumper kablo | 1 | 20cm | ₺3,0 | ₺3,0 |
| Arduino | 1 | Uno | ₺38,0 | ₺38,0 |
| Yazıcı kasası | 1 |  | ₺350,0 | ₺350,0 |
| Genel toplam | ₺497,1 | | | |

**KAYNAKÇA**

-İbrahim Gelegin, Bülent Bolat “AYRIK KELİME TABANLI BİR KONUŞMA TANIMA SİSTEMİYLE BİLGİSAYAR KONTROLÜ”

-İpek Barış, Meltem Erdamar, Emre Sümer, Hamit Erdem “Ses İşaretlerinin Yapay Sinir Ağları ile Tanınması ve Kontrol İşlemleri için Kullanılması”

-Köksal ÖCAL, Yüksek Lisans Tezi ” OTOMATİK KONUŞMA TANIMA ALGORİTMALARININ UYGULAMALARI”