

PROJE ADI: KONUŞMADAN METNE DÖNÜŞÜM

GRUP ADI: MICROMEGA

GRUP ÜYELERİ:

16220054 - KAYAHAN MELİH ALTINÇIBIK

16220014 - BURAK ELHAMAN

16220020 - NURULLAH ALAGÖZ

15220055 - YUNUS YILMAZ

16220062 - MEHMET DAŞÇI

14210034 - AYETULLAH AKDOĞAN

12.12.2019

Bilgisayarların g ndelik yařamın bir par ası olmasıyla birlikte insan-bilgisayar etkileřimi  nem kazanmaya bařlamıřtır. Geleneksel insan-bilgisayar ara y zleri olan klavye, fare gibi aygıtlar yaygın olarak kullanılmalarına raėmen kullanıcı ile bilgisayar arasındaki bilgi ve komut akıřını kısıtlamaktadır. Ayrıca geleneksel ara y zlerin b y k bir kısmı motor yetmezlikleri olan fiziksel engelliler tarafından kullanılamamaktadır. Bu yetersizliklere bir     m olarak bilgisayarın konuřma sinyali ile kontrol edilmesi  nerilebilir. Ancak, leh e ve dilsel farklılıklar gibi nedenlerden dolayı, geniř bir kelime haznesine sahip sistemlerin tasarlanması son derece zordur. Yine de kısıtlı sayıda komutla ger ekleřtirilecek uygulamaların bařarımları tatmin edici d zeylerde olabilmektedir.

Konuřma tanıma, konuřma sinyallerinden otomatik olarak dilsel bilgilerin  ıkarılması olarak tanımlanır. Konuřma tanıma sistemlerinin  alıřma ilkesi kabaca, giriř verisinin daha  nce kaydedilmiř řablonlarla karřılařtırılmasına dayanır. Bu řablonlar fonem d zeyinde ya da kelime d zeyinde d zenlenmiř olabilir. Bu karřılařtırma sonucunda verinin en  ok benzediėi řablon o verinin simgesel g sterimi olarak kabul edilir. Ses sinyallerinin doėrudan karřılařtırılması son derece uzun iřlemler gerektirdiėinden ses sinyalleri  zerinde bilgi miktarını azaltmadan veri boyutunu azaltacak bir  n iřlem yapılması gerekmektedir. Bu  n iřlem  znitelik  ıkarımı olarak adlandırılır. Alınan ses verisinden  znitelik vekt rleri  ıkarılarak elde edilen  znitelikler daha  nce hesaplanarak etiketlenmiř olan řablonlarla karřılařtırılır. Burada elde edilen  znitelik vekt rlerinin ses verilerini en iyi řekilde temsil etmesi, iřlenebilir oran ve b y kl kte olması ve ayırt edici  zelliklere sahip olması gerekmektedir. Literat rde konuřma tanıma uygulamalarında kullanılabilecek bir ok  znitelik tanımlanmıřtır. Bu  zniteliklerden Mel Frekans Kepstral Katsayıları(MFCC) kullanılacaktır.

Konuřma tanıma sistemleri, ayrıık/s rekli, kiřiye baėımlı/kiřiden baėımsız, kelime tabanlı/fonem tabanlı gibi  eřitli kategorilerde sınıflandırılabilirler. Ayrıık konuřma tanıma sistemlerinde, kısa aralıklarla seslendirilen s zc klerin tanınması ama lanırken s rekli konuřma tanıma sistemlerinde ise ara verilmeden seslendirilen s zc klerin tanınması ama lanır. Kiřiye baėlı sistemlerde referans řablonları tek bir kiři i in oluřturulurken, kiřiden baėımsız sistemler tanıma iřlemini konuřmacının kimliėinden baėımsız olarak ger ekleřtirilir. Kelime tabanlı konuřma tanıma sistemleri, tanıma i in  ng r len en k  k birim olarak kelimeyi kabul eder. Fonem tabanlı konuřma tanıma sistemleri ise tanıma i in fonemlerin en k  k birim olarak kabul edildiėi sistemlerdir. B t n konuřma sesleri, farklı frekans deėerlerine sahip sin s dalgalarının doėrusal birleřiminden oluřur. İnsan sesinin frekans deėerleri 300Hz-3300Hz arasında deėiřmektedir. Nyquist Teoremine g re ses frekansının iki katı ve daha b y k  rnekleme frekansı ile etkin bir  rnekleme yapılır.

Projemizde konuřulan kelimelerin Braille alfabesine d n řt r lmesi  zerinde durulmuřtur. Braille alfabesi g rme engellilerin okuyup yazması i in 1821 yılında Louis Braille tarafından geliřtirilmiřtir. Projemizin yazıcı kontrol nde kullanılan yazılımı ve donanımın bir kısmı deėiřtirilerek normal yazılar i in yani yazma engeli olan ve g rebilen insanlar i in de kullanılabilir. Projemizde ayrıık konuřma tanıma kullanılarak kısa aralıklı seslerin s zc klerinin tanınması, kelime tabanlı olarak ve kullanacaėımız yapay zeka sayesinde kiřiiden baėımsız olarak kullanılması ama lanmaktadır.

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 tanıma - 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 Altınçılık
3	Yazıcı bölümünün yapılması ve yazılımının yazılması	Nurullah Alagöz Kayahan Melih Altınçılık 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 mikrodnetleyici kullanmayı düşünmüştük. Bunun için mikrodnetleyicilerin mikrofon ve bluetooth ile etkileşimlerini araştırıp kullanmaya çalıştık. Ancak sonradan projemiz için hafıza ve ram ihtiyacının mikrodnetleyicilerin çoğunun karşılamayacağını farkettilik. Bu nedenle Matlab kullanmaya karar verdik. Matlab kullanmamızın bir diğler 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 değıştiriyoruz. 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. Çerçeveleme 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.

$$w(n) = \begin{cases} 0.54 - 0.46 \cos(2\pi \frac{n}{N-1}) & ; 0 \leq n \leq N-1 \\ 0 & ; \text{diğler} \end{cases}$$

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.

$$y(n) = x(n) - 0.95x(n - 1)$$

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

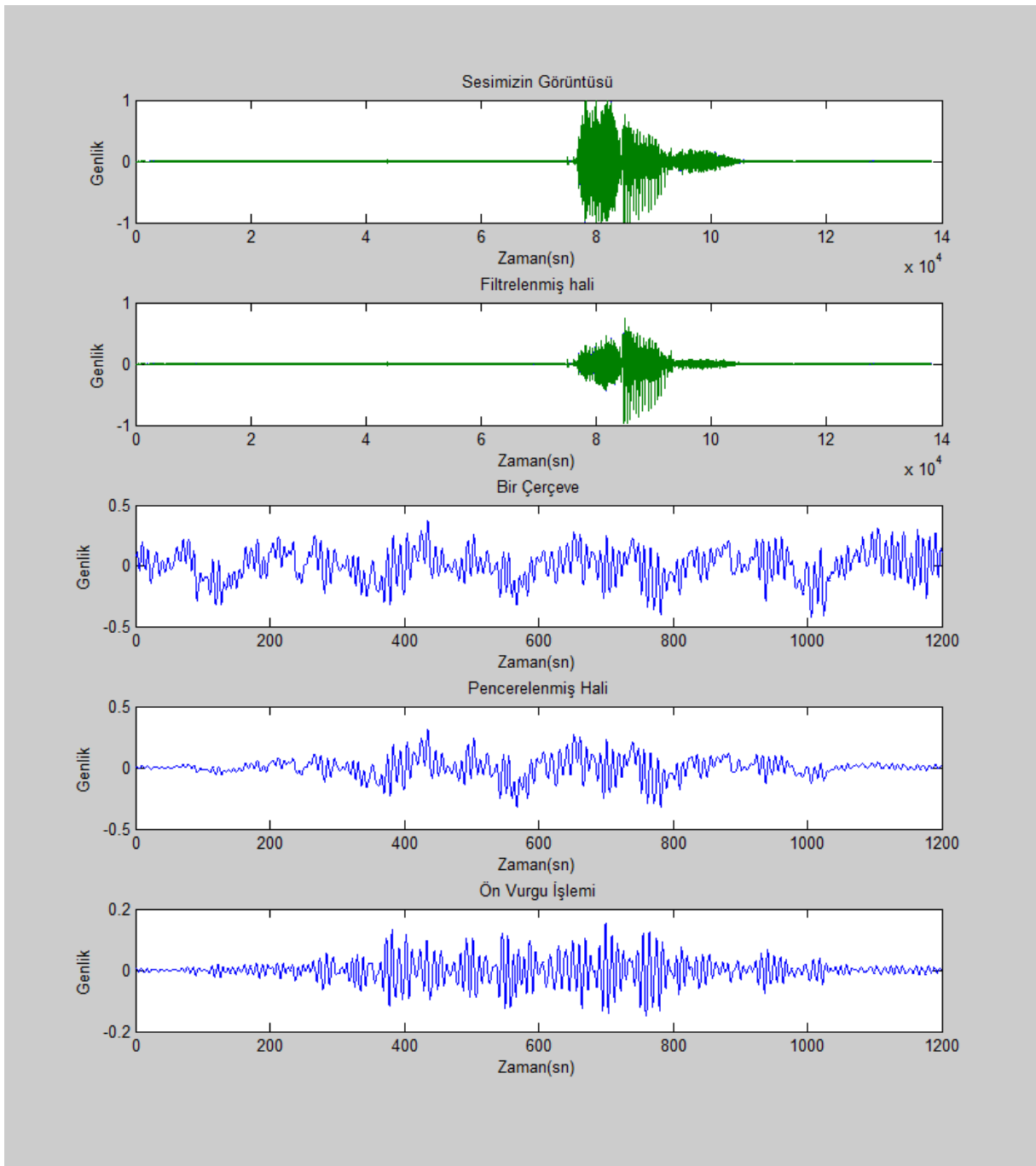
subplot(5,1,1)
plot(y)
xlabel('Zaman(sn) ');ylabel('Genlik');title('Sesimizin
Görüntüsü');
subplot(5,1,2)
plot(a)
xlabel('Zaman(sn) ');ylabel('Genlik');title('Filtrelen
miş hali');
subplot(5,1,3)
plot(fr(:,1:135))
xlabel('Zaman(sn) ');ylabel('Genlik');title('Bir
Çerçeve');
subplot(5,1,4)
```

```

plot(fr_win(135,:))
xlabel('Zaman(sn)');ylabel('Genlik');title('Pencerele
nmiř Hali');
subplot(5,1,5)
plot(e(135,:))
xlabel('Zaman(sn)');ylabel('Genlik');title('Ön Vurgu
İřlemi');

```

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”