

SAYISAL İŞARET İŞLEME TASARIM ve UYGULAMALARI PROJE ÖNERİ FORMU

1. PROJE BAŞLIĞI:

Dile Duyarlı Kelime Tanıma

2. PROJE TAKIMI:

1 : Nurefşan Sertbaş
2 : Meryem Meray Yağmur
3 : Shamyra Zyrriyev
Danışman : Bahri Abacı

3. KONU:

Telaffuz edilen kelimenin hangi dilde olduğunun algılanması ve Türkçe karşılığının bulunması. Günümüzde üzerinde sıkça çalışılan popüler bir konu olmakla birlikte projenin altında yatan temel fikir 1900 lü yıllara kadar uzanır. Ancak zorluğu ve ses sinyallerindeki çeşitlilik gibi sebeplerden ötürü günlük hayata girmesi zaman almıştır. Modern donanımlar ve yeni nesil algoritmalar ile hayat bulması yakın zamanda Google tarafından mümkün olmuştur (2006). Hayata geçirilen projede 64 dil desteği sağlanmaktadır.

4. HEDEFLER:

Bu projede amaçlanan ortam gürültüsünün mevcut olduğu bir durumda kişinin konuşmasının donanım yardımıyla alınması, MATLAB kullanılarak işlenerek tanınması daha sonra da oluşturulan 'örnek veri tabanından' kelimenin tanımlı olduğu dilin tespit edilmesidir.

5. KAPSAM:

Bu projenin gerçekleşmesinde özellikle de elde edilen çıktıların belli bir doğruluk seviyesinin üstünde kalması adına 4 ana prensip göz önünde bulundurulmaktadır. Bunlardan ilki 'acoustic models' dediğimiz konuşmanın neye benzediği, ikincisi 'pronunciations' denilen kelimenin nasıl telaffuz edildiği, üçüncüsü 'vocabulary' denilen ilgili kelimeyle anlatılmak istenen ve son olarak da 'language model' ile kelimelerin nasıl bir araya geldiğidir. Her dilde telaffuz, o telaffuza denk gelen sözcük ve kelimelerin bir araya geliş şekilleri, cümle yapıları farklı olduğundan diller arası büyük farklılıklar mevcuttur. Bu sebeple projenin cümle yerine kelime bazında kalması ve belli dillere destek vermesi planlanmıştır. Yapacağımız projede Türkçe, İngilizce, Rusça, Türkmençe ve Kırgızca gibi dillerin bir kısmına destek verilmesiyle Google tarafından gerçekleştirilen örnek modelden nispeten daha küçük bir arşiv oluşturulması planlanmaktadır. Bahsedilen arşiv teknik olarak bir veritabanı olabileceği gibi bir dosya da olabilir. Ayrıca tanımlı olacak kelime sayısı da sınırlı tutulacaktır.

6. YÖNTEM:

Planlanan projenin gerçekleşmesi sırasında baştan sona kadar sayısal işaret işlemeye dair mevcut tekniklerin implemente edilmesi sağlanacak ve gereken yerlerde yeni algoritmalar kurulacaktır. Projenin çeşitli aşamalarında başvurulması olası tekniklerin bir kısmı aşağıda listelenmiştir:

- 'Fourier Dönüşümü' sayesinde sinyal hakkında karakteristik bilgiler elde edilebilir.
- Yine ses tanıma projelerinde sıklıkla yer alan 'Hidden Markov modeli'. Sınırlı dağarcıklı ayırık sözcük tanıma olarak tanımlanan projemizde her sözcüğe ait bir HMM modeli oluşturulup, mevcut sesin bu modellerle işlenmesi esas alınmaktadır. Modelden elde edilen sonuçlar en yakın kategoriye koyularak uygun karşılık seçilir.
- 'Dinamik Zaman Eşleştirme' de ise kişilerin ifadeleri seslendirme süreleri sıkıştırılarak ya da genişletilerek referanslarla karşılaştırılmaları ilkesi kullanılmaktadır. Aynı kişinin bile aynı sözcüğü iki farklı zamanda farklı uzunluklarda söyleyebileceği göz önüne alındığında kullanımı daha iyi anlaşılacaktır. Amaç referansla aynı aralıkta karşılaştırma yapabilmektir.
- Bazı ses ile tanımatanım projelerinde ise *Mel Frequency Cepstral Coefficients (MFCC)* kullanılır.

7. KATKILAR:

- Bu proje ve kullanımıyla gerçekleştirilecek diğer projelerin toplumun her kesimine hitap etmesi ve fayda sağlaması beklenmektedir.
- Bu projenin uyarlanmasıyla görme engellilere yardımcı olacak yeni projeler geliştirilebilir.
- Ses kontrollü donanımların kullanımında artış söz konusu olabilir.İnsan-makine iletişimde tuşları aradan kaldırır.Bu sayede doğrudan iletişime olanak tanır.Dolayısıyla hız artarken maliyette düşüşler gözlenir.
- Bilimsel olarak insan konuşmasının dijital ortamda işlenmesi farklı uygulamaların geliştirilmesi için teşvik edici olur.
- Çağrı merkezleri vs gibi telefon işlemlerinde(interaktif sesli cevap sistemleri) müşteri hizmetlerinin kalitesini arttırır.

8. PROJE YÖNETİMİ:

Ekip üyelerinin tamamı projenin her aşamasından doğrudan sorumlu olacak ancak raporlama sırasında görev paylaşımı yapılacaktır.

9. ÇALIŞMA PLANI:

1. 03.03.2015(5.hafta)-17.03.2015(7.hafta) :Ses işlemeye yönelik literatür taramasının yapılması
2. 17.03.2015(7.hafta)-21.04.2015(12.hafta) :Literatüre ek olarak programlamaya başlanması
3. 21.04.2015(12.hafta)-05.05.2015(14.hafta) :Sonuçların raporlanması, gerekli görsel arayüzlerin tasarlanması, sunuma hazırlık

10. PERFORMANS KRİTERLERİ:

Bu projenin başarılı sonuçlanabilmesi,öncelikle projede hedeflenen sonuca ulaşabilmek için planlanan süreçlerin zamanında,planlı,programlı ve arzu edilen kalitede bir çıktı ile teslim edilebilmesine bağlıdır.Proje başlangıcı ile projenin teslim edildiği tarih arasında proje hedef ve çıktılarındaki büyük değişiklikler yapılmayarak başarılı bir hedefe ulaşmak planlanmıştır.Performans ölçütlerinde doğruluk ve hız esas alınmaktadır.

Bu projenin gerçekleştirilmesini engelleyecek herhangi bir durum söz konusu değildir.Ancak ve ancak ihtimal üzerine, input olarak alınan ses dosyasının test aşamasında beklenen sonuçları yansıtmaması veya farklı alfabeler söz konusu olduğundan yazıya dökülmesi sırasında çıkabilecek birtakım problemler söz konusu olabilir.Böyle bir durum da dahi ses dosyasını belli bir dil için yazıya döken algoritma geliştirmek bile başarı sayılabilir.

Zamanın kısıtlı oluşu sebebiyle yeterli miktarda dile destek verilememesi ve kelime dağarcığının kısıtlı kalması söz konusu olabilir.Bu durum B planı olarak gösterilebilir.

11. BÜTÇE:

Proje yazılım tabanlı olduğundan dolayı herhangi bir bütçeye ihtiyaç duyulmayacaktır.

Ekler:

1. Her grup 03.03.2015 saat 13:00 a kadar proje öneri formunu ninovaya yüklemeli ve ders başlangıcında çıktı olarak teslim edilmelidir.
2. Öneri formuna yapılmak istenen ekler formata uygun açılan kutular içerisinde yapılabilecektir.
3. Projenin tamamlanması sonucu beklenen not aşağıda belirtilmelidir.

Beklenen/Hak edilen Not: BA(min)