

Mobil Ortamda Konuşma Tanıma Teknolojisinin İncelenmesi

Nihal Yağmur Aydın

Boğaziçi Üniversitesi

Taslak

- Giriş
- Önceki Çalışmalar
- Google Speech API ile masaüstü uygulaması
- Android Speech API ile mobil uygulama
- Tartışma ve İleri Çalışmalar

Giriş

- Konuşma tanıma teknolojisinin mobil ve masaüstü ortamlarda test edilmesi amacıyla bu çalışma yapılmıştır.
- Mobil ortamda sunucuya gönderilen ses girdisinin yazı biçiminde dönüşümü ile gelen sonuçlar değerlendirilmiştir.

Önceki Çalışmalar

- Google Speech API: anahtarla kullanılır.
Http post isteği ile sunucudan bilgi döner.
- Kaldi: C++ dilinde yazılmış, offline olarak konuşma tanıma işlemini gerçekleştiren araç.
- CloudASR : Kaldi'nin Online hale getirilmesi sonucu yaratılan konuşma tanıma sistemi.

Google Speech API ile masaüstü uygulaması

- (*.wav) formatındaki dosya sunucuya gönderilir.
- Bu işlemin sonucunda, Google sunucusundan dönen çıktı aşağıdaki gibidir:

```
{"result":[]}
```

```
{ "result":[
```

```
  { "alternative":[
```

```
    { "transcript":"umbrella",
```

```
      "confidence":0.97527361 }, {
```

```
      "transcript":"unbrella" }, {
```

```
      "transcript":"umbrellas" }, {
```

```
      "transcript":"Tom Brella" } ], "final":true } ], "result_index":0 }
```

Android Speech API ile Mobil Uygulaması

- Örnek (kelime bazında, yerel telefon dili Türkçe)
“elephant” kelimesi ile test edildiğinde,

Program çıktısı:

“1. Elephant, 2.telefon, 3.telefonun, 4.telefonu, 5. Elif”

Android Speech API ile Mobil Uygulaması

- Örnek (kelime bazında, yerel telefon dili İngilizce)

“elephant” kelimesi ile test edildiğinde,

Program çıktısı:

“1. elephant, 2. elephants, 3. elefante, 4. indifference, 5. Elephant”

Tartışma ve İleri Çalışmalar

- Android Speech API ve Google Speech API, internet bağlantısı gerektirmektedir. Bu da, telefonun güç kaynağını kullandığı için, performans düşüklüğüne neden olmaktadır.
- Ortamda gürültü olduğunda Google sunucusundan gelen sonuç gerçekçi olmamaktadır.
- Mobil uygulamaya filtreleme işlemi getirildiğinde performans kaybı olur, çünkü androidin sınırlı bir RAM'i vardır.
- Sunucu bazlı geliştirmeler yapılabilir.
- Engelliler için uygulamalar geliştirilebilir.

Kaynaklar

[1] CloudASR: Platform and Service, by Ondřej Klejch, Ondřej Plátek, Lukáš Zilka, Filip Jurčiček (2015), Springer Publishing. Chapter: Text, Speech, and Dialogue, Volume 9302 of the series Lecture Notes in Computer Science pp 334-341. 033-6

[2] <http://developer.android.com/reference/android/speech/SpeechRecognizer.html>

[3] Chang Y., Hung S , Wang N J.C., Lin B. (2011), *CSR: A Cloud-Assisted Speech Recognition Service for Personal Mobile Device* , IEEE International Conference on Parallel Processing, pp. 305 - 314

[4] <https://developers.google.com/web/updates/2013/01/Voice-Driven-Web-Apps-Introduction-to-the-Web-Speech-API>

[5] <https://gist.github.com/alotaiba/1730160>

[6] <http://developer.android.com/reference/android/speech/tts/TextToSpeech.html>

[7] <http://kaldi.sourceforge.net/>

[8] <http://www.cloudasr.com/>