
Trabalhando com reconhecimento de voz

Henrique Silva Anastácio
Ítalo Dell'Areti
Rita Rezende
Artur Parreiras

O que é ?

- O reconhecimento da fala permite que computadores, celulares ou tablets equipados com microfones, interpretem a fala humana para transcrição ou como método de comando por voz.
- Podem ser classificados por requererem (ou não), que o usuário treine o sistema a reconhecer seus padrões particulares de fala.
 - Sistemas que requerem pouco treinamento, podem capturar continuamente a fala com um amplo vocabulário em ritmo normal, com precisão de cerca de 98%.
 - Enquanto sistemas que não requerem treinamento podem reconhecer um número pequeno de palavras, como os dez dígitos do sistema decimal.
 - Tais sistemas são populares por direcionar chamadas telefônicas recebidas, em grandes organizações, aos seus destinos.
- Disponíveis desde os anos 90.

Como funciona?

- É utilizado um conversor analógico-digital que capta as vibrações criadas pela sua voz e converte essas ondas em dados digitais.
- O som digitalizado é filtrado para separá-lo de ruídos e interferências. (O som pode necessitar ser sincronizado, pois as pessoas não costumam utilizar o mesmo tom e nem sempre falam na mesma velocidade).

Como funciona?

- Em seguida, o programa compara os sons captados com fonemas conhecidos e presentes em seu banco de dados, que correspondam ao idioma que o locutor tenha falado.
- Por último, o sistema analisa o resultado e o compara com palavras e frases conhecidas, assim, ele identifica o que seu usuário disse e converte para a funcionalidade desejada (texto em uma planilha, um comando, etc.).
- Essa digitalização é separada em frações ainda menores (sons fonéticos não maiores do que uma sílaba).

Exemplo

```
// Importando
import android.speech.RecognizerIntent;

public class MainActivity extends AppCompatActivity {

    // ID
    private final int ID_TEXTO_VOZ = 1;

    Button btnfala;
    TextView txtcapturado;

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.activity_main);

        btnfala = (Button) findViewById(R.id.botao_fala);
        txtcapturado = (TextView) findViewById(R.id.texto_falado);

        btnfala.setOnClickListener(new View.OnClickListener() {
            @Override
            public void onClick(View v) {

                // Reconhecimento de voz
                Intent intentVoz = new Intent(RecognizerIntent.ACTION_RECOGNIZE_SPEECH);

                // Setar idioma e modo de fala livre
                intentVoz.putExtra(RecognizerIntent.EXTRA_LANGUAGE_MODEL, RecognizerIntent.LANGUAGE_MODEL_FREE_FORM);

                // Pega a lingua padrão do celular e reconhece o que é dito através dela
                intentVoz.putExtra(RecognizerIntent.EXTRA_LANGUAGE, Locale.getDefault());

                // Mensagem Prompt
                intentVoz.putExtra(RecognizerIntent.EXTRA_PROMPT, "Fale o que desejar");

                try{
                    startActivityForResult(intentVoz, ID_TEXTO_VOZ);
                }

                catch (ActivityNotFoundException a){
                    Toast.makeText(getApplicationContext(), "Seu celular não suporta comando de voz",
                        Toast.LENGTH_SHORT).show();
                }
            }
        });
    }
}
```

Exemplo

```
@Override
protected void onActivityResult(int id,int resultCodeID,Intent dados){
    super.onActivityResult(id,resultCodeID,dados);
    switch (id){

        case ID_TEXTO_VOZ:
            // Verifica o que foi falado
            if (resultCodeID == RESULT_OK && null != dados){

                //Passa para o array
                ArrayList<String> result = dados.getStringArrayListExtra(RecognizerIntent.EXTRA_RESULTS);
                String ditado = result.get(0);

                //Mostra texto via Toast
                Toast.makeText(getApplicationContext(),ditado,Toast.LENGTH_SHORT).show();

                // Mostra o que foi falado no textView
                txtcapturado.setText(ditado);
            }
        break;
    }
}
```

Exemplo - 2

Inicialização

- Para começar usando a biblioteca, você tem que colocar isso no seu Application subclass:

```
public class Initializer extends Application {  
  
    @Override  
    public void onCreate() {  
        super.onCreate();  
  
        Speech.init(this);  
    }  
}
```

Exemplo - 2

Logging

- Por padrão, a biblioteca logging está desativada. Você pode ativar o debug log implementando:

```
Logger.setLogLevel(LogLevel.DEBUG);
```

- Onde você quiser no seu código você pode ajustar o nível de detalhe do DEBUG para OFF.
- A biblioteca logger usa *android.util.Log* por padrão, assim, você obterá a saída no LogCat. Se você quiser redirecionar os logs para saídas diferentes ou usar um diferente, você pode fornecer sua própria implementação delegada da seguinte forma:

Exemplo - 2

```
- Logger.setLoggerDelegate(new Logger.LoggerDelegate() {  
    @Override  
    public void error(String tag, String message) {  
        //your own implementation here  
    }  
  
    @Override  
    public void error(String tag, String message, Throwable exception) {  
        //your own implementation here  
    }  
  
    @Override  
    public void debug(String tag, String message) {  
        //your own implementation here  
    }  
  
    @Override  
    public void info(String tag, String message) {  
        //your own implementation here  
    }  
});
```

Exemplo - 2

```
@Override
public void onSpeechPartialResults(List<String> results) {
    StringBuilder str = new StringBuilder();
    for (String res : results) {
        str.append(res).append(" ");
    }

    Log.i("speech", "partial result: " + str.toString().trim());
}

@Override
public void onSpeechResult(String result) {
    Log.i("speech", "result: " + result);
}
});
```

Exemplo - 2

```
} catch (SpeechRecognitionNotAvailable exc) {  
    Log.e("speech", "Speech recognition is not available on this device!");  
    // You can prompt the user if he wants to install Google App to have  
    // speech recognition, and then you can simply call:  
    //  
    // SpeechUtil.redirectUserToGoogleAppOnPlayStore(this);  
    //  
    // to redirect the user to the Google App page on Play Store  
}
```

Exemplo - 2

Exibir animação de progresso

Adicione isso no layout:

```
<LinearLayout
    android:orientation="vertical"
    android:layout_width="wrap_content"
    android:layout_height="wrap_content"
    android:id="@+id/linearLayout">

    <net.gotev.speech.ui.SpeechProgressView
        android:id="@+id/progress"
        android:layout_width="120dp"
        android:layout_height="150dp"/>

</LinearLayout>
```

Problemas

- Diferenças entre os interlocutores são frequentemente grandes e dificultam. Não está claro quais características da fala são independentes do falante.
- A interpretação de vários fonemas, palavras e frases é sensível ao contexto. Por exemplo: os fonemas são geralmente mais curtos em palavras longas do que em palavras pequenas.

Problemas

- A entonação e o timbre da fala podem mudar completamente a interpretação de uma palavra ou frase. Por exemplo: "Vai!", "Vai?".
- Palavras e frases podem ter várias interpretações válidas de modo que o falante deixe a escolha da correta para o ouvinte.
- A linguagem escrita precisa de pontuação de acordo com regras estritas, que não estão fortemente presentes na fala e são difíceis de inferir sem conhecer o significado.

Futuro



Referências

<https://developer.android.com/reference/android/speech/RecognizerIntent>

<https://developer.android.com/reference/android/speech/SpeechRecognizer>

<https://github.com/zagum/SpeechRecognitionView>

<https://www.tecmundo.com.br/curiosidade/3144-como-funciona-o-reconhecimento-de-voz-.htm>

https://pt.wikipedia.org/wiki/Reconhecimento_de_fala