

A "voz"... é a forma mais natural de comunicar.



Reconhecedor de comandos robusto embebido

Luis Vieira de Sá

Sumário

- Objectivos
- Tecnologia
- Base de Dados
- Integração
- Demo

Equipa

- 3 PhDs
- 2 MScs (1)
- 5 Engenheiros (3)



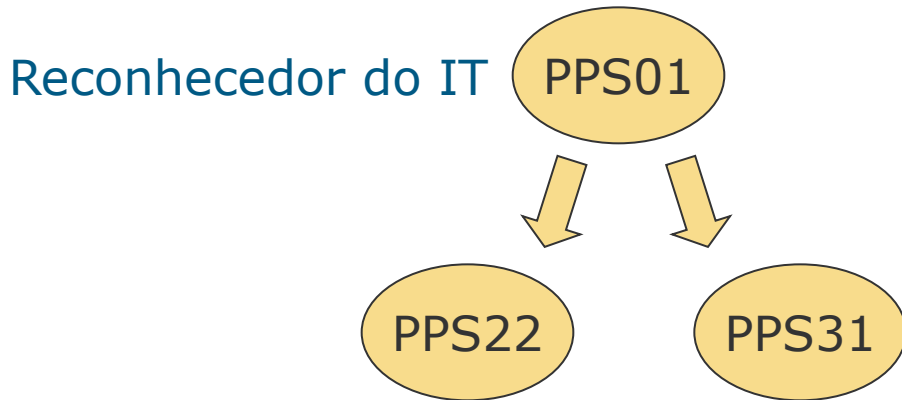
Objectivos - Gerais

- Comandos ligados
- Independência do locutor
- Vocabulário reduzido (<250)
- Robustez a ruído
- Medidas de confiança
- Hardware específico



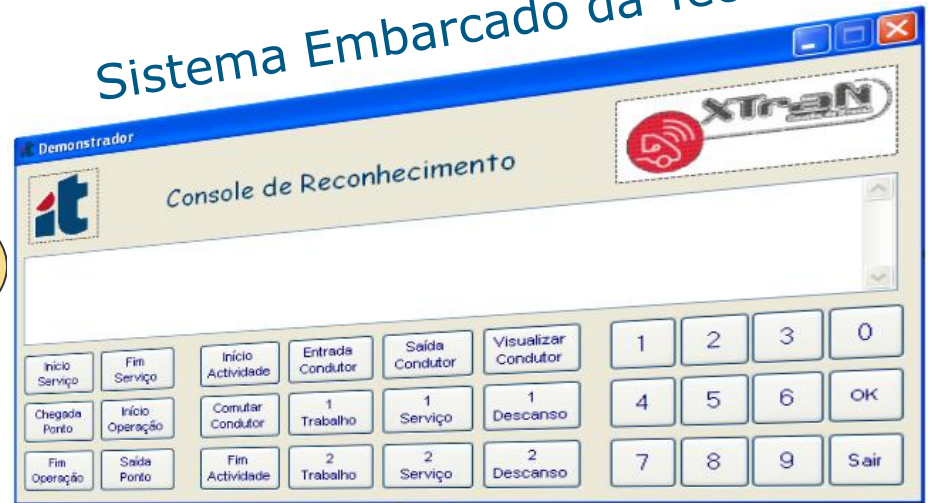


Objectivos - Demonstradores



Sistema Controlador

Sistema Embarcado da Tecmic

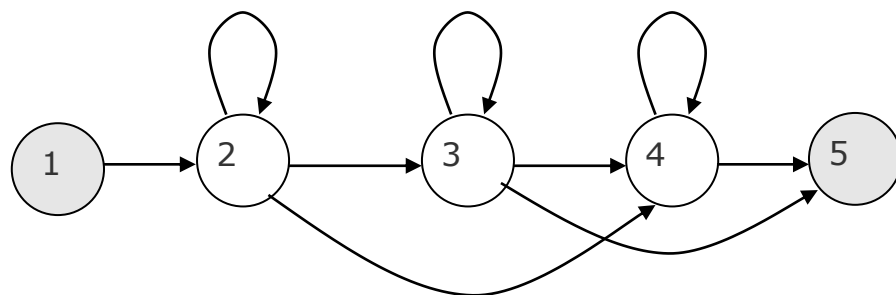


O sistema deverá funcionar em ambientes acústicos "hostis", nomeadamente no interior de automóveis ou em ambientes industriais.



Tecnologia - Modelos

- CDHMM - Continuous density hidden Markov models
- MFCC – Mel frequency cepstrum coefficients

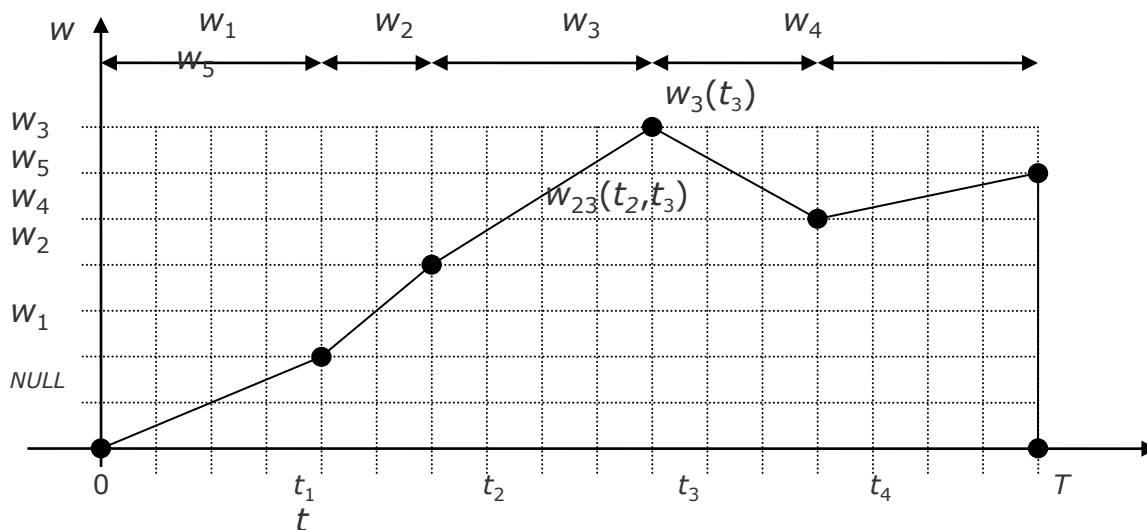


- 250 modelos
- 3 estados/fonema
- 39 parâmetros MFCC
- 8 misturas



Tecnologia - Descodificador

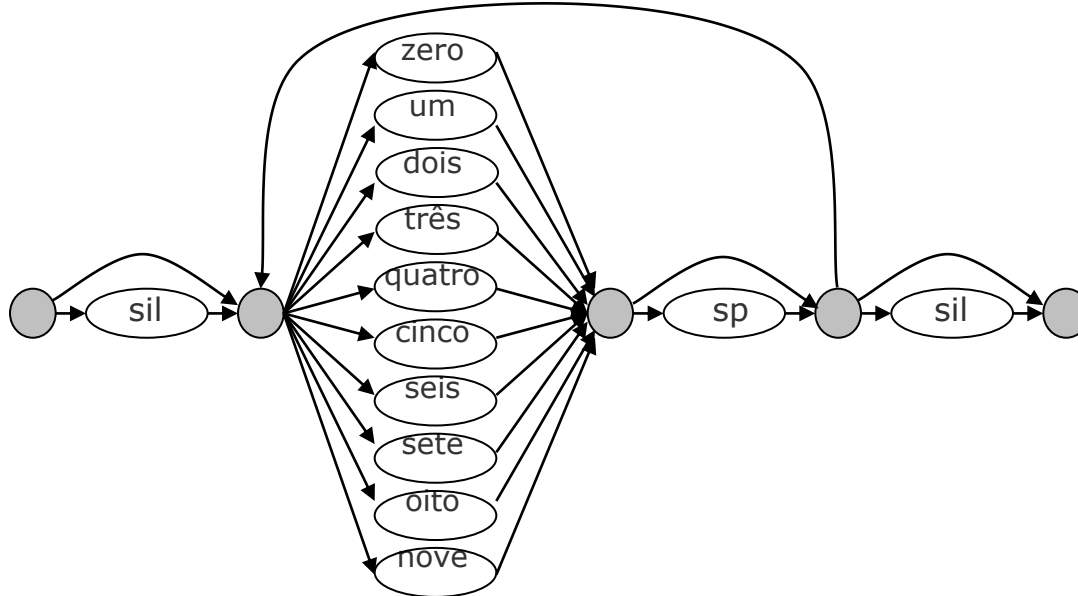
- Descodificador *token-passing*





Tecnologia - Gramáticas

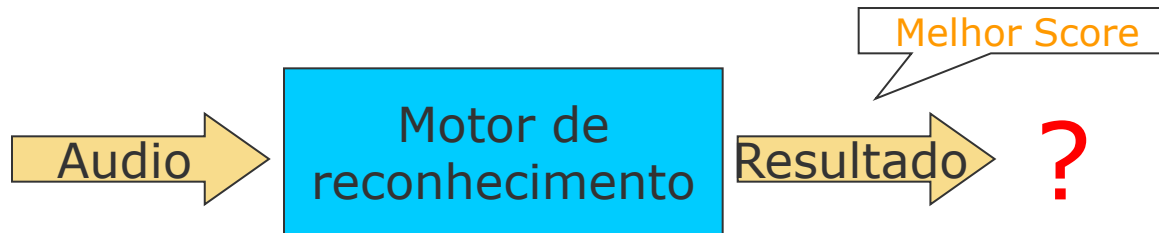
- <http://www.w3.org/TR/speech-grammar/>





Tecnologia – Medidas de Confiança (1)

- Munir o resultado do reconhecimento com um valor que traduz a sua confiança.

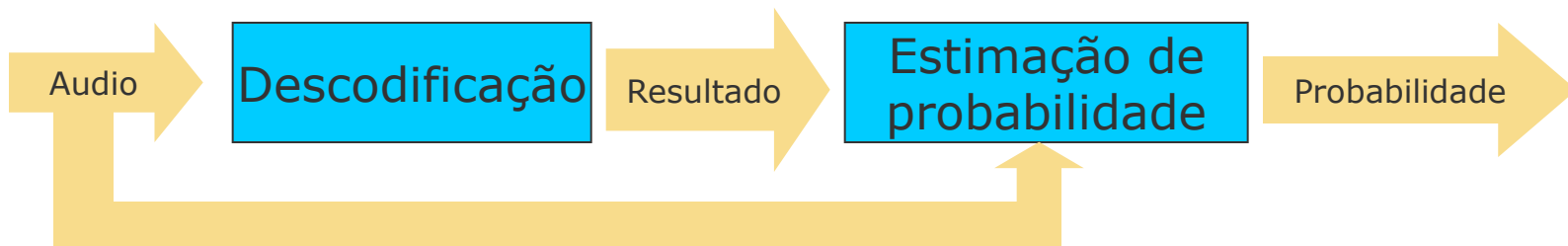


- O classificador só disponibiliza o melhor melhor score obtido com o algoritmo de viterbi (likeH0).

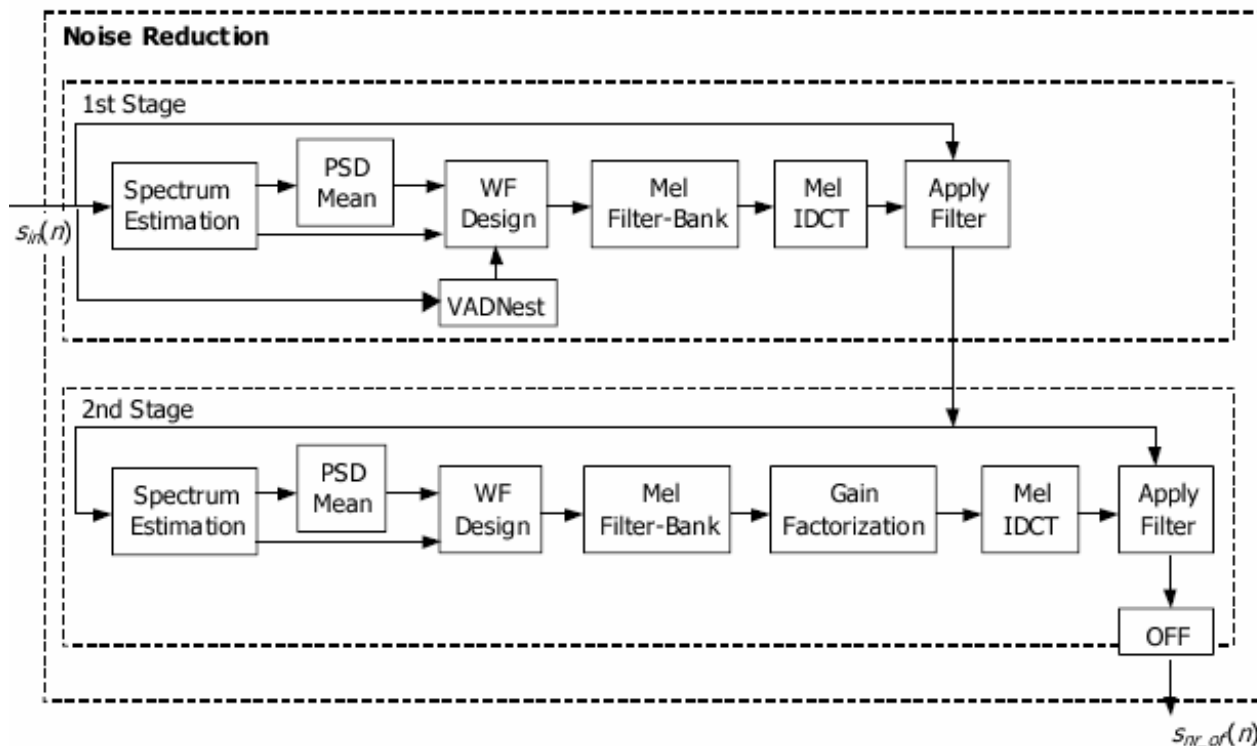


Tecnologia – Medidas de Confiança (2)

Usa-se um modelo de “lixo” (filler) para estimar a probabilidade dos resultados estarem certos a partir do LLR (log likelihood ratio)



Tecnologia – Redução de ruído





Base de Dados

- 4 semanas de gravações
- Sessões de 30 minutos
(2 a 4 canais)
- 239 sessões (locutores)
- 67,7 GB
- 16 DVDs gravados



Administração das Sessões de Gravação

Click sobre o link de um dia para ver o plano de sessões para esse dia. Click sobre o link de uma sessão para ver ou editar essa sessão.

	14	15	16	17	18	21	22	23	24	25	28	29	30	31	01	04	05	06	07	08
	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Maio	Junho	Junho	Junho	Junho	Junho	Junho
09h00	S001	S023	S045	S067	S089	S111	S133	S155	S177	S199	S221	S243	S265	S287	S309	S331	S353	S375	S397	S419
09h30	S002	S024	S046	S068	S090	S112	S134	S156	S178	S200	S222	S244	S266	S288	S310	S332	S354	S376	S398	S420
10h00	S003	S025	S047	S069	S091	S113	S135	S157	S179	S201	S223	S245	S267	S289	S311	S333	S355	S377	S399	S421
10h30	S004	S026	S048	S070	S092	S114	S136	S158	S180	S202	S224	S246	S268	S290	S312	S334	S356	S378	S400	S422
11h00	S005	S027	S049	S071	S093	S115	S137	S159	S181	S203	S225	S247	S269	S291	S313	S335	S357	S379	S401	S423
11h30	S006	S028	S050	S072	S094	S116	S138	S160	S182	S204	S226	S248	S270	S292	S314	S336	S358	S380	S402	S424
12h00	S007	S029	S051	S073	S095	S117	S139	S161	S183	S205	S227	S249	S271	S293	S315	S337	S359	S381	S403	S425
12h30	S008	S030	S052	S074	S096	S118	S140	S162	S184	S206	S228	S250	S272	S294	S316	S338	S360	S382	S404	S426
13h00	S009	S031	S053	S075	S097	S119	S141	S163	S185	S207	S229	S251	S273	S295	S317	S339	S361	S383	S405	S427
13h30	S010	S032	S054	S076	S098	S120	S142	S164	S186	S208	S230	S252	S274	S296	S318	S340	S362	S384	S406	S428
14h00	S011	S033	S055	S077	S099	S121	S143	S165	S187	S209	S231	S253	S275	S297	S319	S341	S363	S385	S407	S429
14h30	S012	S034	S056	S078	S100	S122	S144	S166	S188	S210	S232	S254	S276	S298	S320	S342	S364	S386	S408	S430
15h00	S013	S035	S057	S079	S101	S123	S145	S167	S189	S211	S233	S255	S277	S299	S321	S343	S365	S387	S409	S431
15h30	S014	S036	S058	S080	S102	S124	S146	S168	S190	S212	S234	S256	S278	S300	S322	S344	S366	S388	S410	S432
16h00	S015	S037	S059	S081	S103	S125	S147	S169	S191	S213	S235	S257	S279	S301	S323	S345	S367	S389	S411	S433
16h30	S016	S038	S060	S082	S104	S126	S148	S170	S192	S214	S236	S258	S280	S302	S324	S346	S368	S390	S412	S434
17h00	S017	S039	S061	S083	S105	S127	S149	S171	S193	S215	S237	S259	S281	S303	S325	S347	S369	S391	S413	S435
17h30	S018	S040	S062	S084	S106	S128	S150	S172	S194	S216	S238	S260	S282	S304	S326	S348	S370	S392	S414	S436
18h00	S019	S041	S063	S085	S107	S129	S151	S173	S195	S217	S239	S261	S283	S305	S327	S349	S371	S393	S415	S437
18h30	S020	S042	S064	S086	S108	S130	S152	S174	S196	S218	S240	S262	S284	S306	S328	S350	S372	S394	S416	S438
19h00	S021	S043	S065	S087	S109	S131	S153	S175	S197	S219	S241	S263	S285	S307	S329	S351	S373	S395	S417	S439
19h30	S022	S044	S066	S088	S110	S132	S154	S176	S198	S220	S242	S264	S286	S308	S330	S352	S374	S396	S418	S440

Treino: 103.001

Teste: 27.550

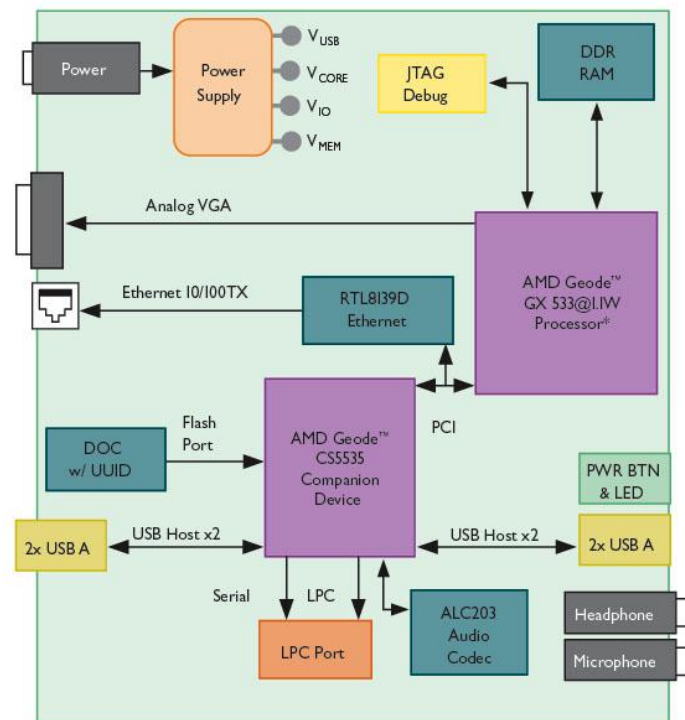
Desenvolvimento: 6.680

Total: 137.439



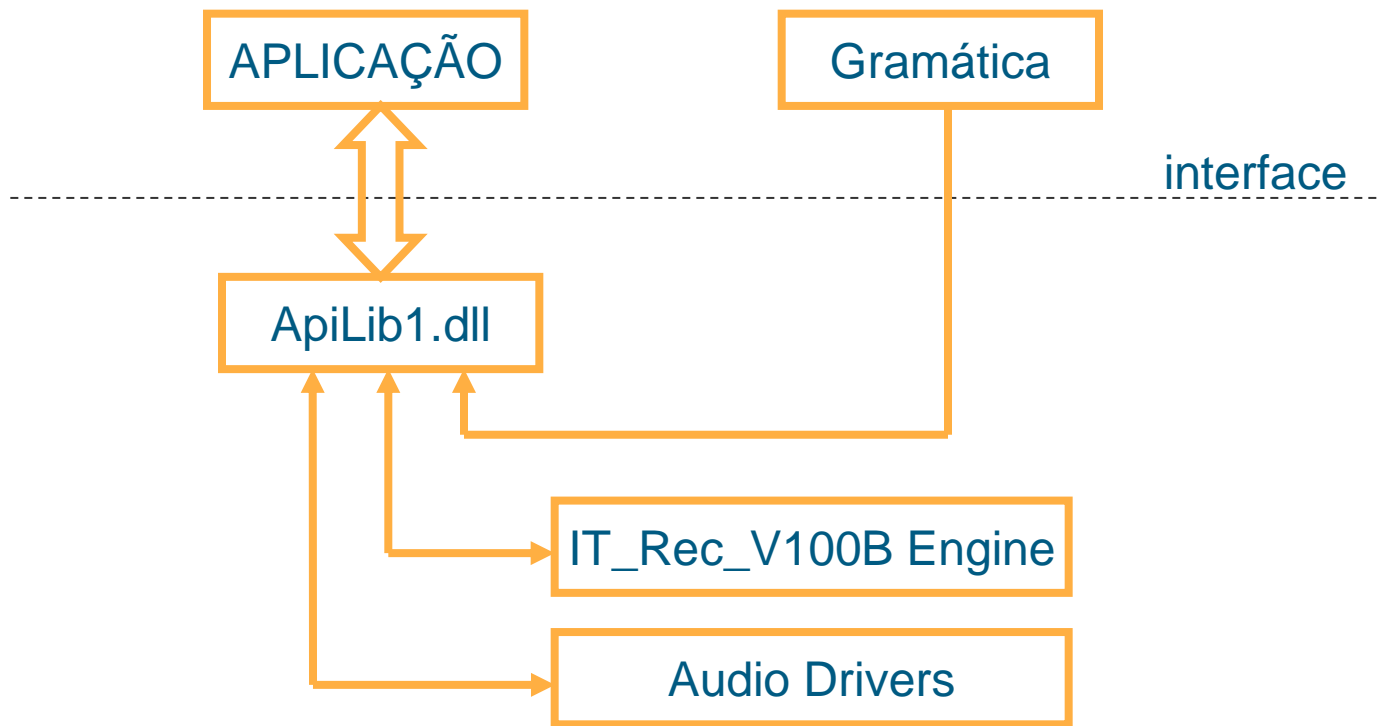
Integração - Hardware

- AMD Geode





Integração – API (1)





Integração – API (2)

```
Imports apiLibv1

Public Class Form1
    Dim config As New SRConfig
    Dim engine As New SREngine

    Private Sub Form1_Load(ByVal sender As Object, _
        ByVal e As System.EventArgs) Handles Me.Load
        config.engineDescription = "IT_Rec_V100A Engine"
        config.fichGram = "digitos.xml"
        AddHandler engine.OnSREcognition, AddressOf On_Recognition
        engine.interest = SREngine.EventInterest.SREcognition
        engine.InitSREngine(config)
        engine.StartRecognition()

    End Sub

    Sub On_Recognition(ByVal sender As Object, _
        ByVal result As SREngine.SREcognitionEventArgs)
        Label1.Text = result.utterance

    End Sub
End Class
```



DEMO

