

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Método de Decisão de Reconhecimento de Comando de Voz para Sistema de Condução de Cadeira de Rodas

Ariana Moura Cabral ¹

Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia

1 Introdução

A Tecnologia Assistiva corresponde a um campo da Engenharia em grande desenvolvimento, já que versa pela inclusão social, isto é, busca a integração, a independência e a melhor qualidade de vida daqueles que apresentam algum tipo de deficiência. Nesse cenário, o desenvolvimento de tecnologias assistivas baseadas no reconhecimento de voz configura-se como uma excelente interface homem-máquina [1], principalmente, quando se pensa em pacientes cadeirantes incapazes de se locomover de modo direto ou com capacidade de movimento dos membros restrita.

À luz disso, o presente trabalho tem como principal objetivo apresentar a implementação de um método de decisão voltado para o reconhecimento de comandos de voz de palavras isoladas, utilizando-se o conceito de distância euclidiana, para um sistema de condução de cadeira de rodas. Nesse sentido, tal processo de reconhecimento é realizado por meio do ambiente de desenvolvimento integrado para linguagem R, o *RStudio*[®].

2 Metodologia

Não somos capazes de pronunciar um palavra ou locução sempre da mesma forma, além disso é notória a distinção entre diferentes fonéticas. Tal fato ocorre por causa da complexidade do sistema biológico responsável pelo sinal de voz e da influência de diversas variáveis (como o ruído) em seu processo de geração. Assim, o reconhecimento de voz nada mais é do que um problema de padrões, configurado através de uma sequência de parâmetros que caracterizam o sinal de voz. Portanto, o ponto crucial da problemática reside na identificação correta dos comandos, mais especificamente, consiste em encontrar padrões de repetições em cada um dos comandos para os diferentes sinais de entrada do sistema.

Logo, ao se objetivar o reconhecimento de comandos isolados (no caso, dos comandos para condução da cadeira de rodas: *direita*, *esquerda*, *para frente*, *para trás*, *parar*, *rotacionar*) na implementação computacional, o conceito de *distância euclidiana* apresenta-se como uma ferramenta extremamente poderosa na investigação de padrões [2], isto é, na

¹arianacabral57@gmail.com

verificação de dissimilaridade entre elementos, já que se almeja estabelecer a natureza dos dados dos sinais gerados para agrupá-los. Assim, ao considerar a distância d entre dois conjuntos de dados representantes dos picos característicos de cada sinal, ou seja,

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_{20} - y_{20})^2},$$

em que x_1, \dots, x_{20} e y_1, \dots, y_{20} correspondem às coordenadas dos 20 maiores picos representantes de dois determinados sinais (referentes ao mesmo comando ou não), consegue-se identificar quais sinais apresentam uma significativa distinção com a iteração desse método de comparação.

3 Resultados e Discussões

Os resultados da implementação em R dos sinais gerados e janelados no software Audacity podem ser visualizados através da tabela, Figura 1, que ilustra a matriz de (dis)similaridade.

	dE1.1	dE1.2	dE1.3	dE1.4	dE1.5	dE2.1	dE2.2	dE2.3	dE2.4	dE2.5	dE3.1	dE3.2	dE3.3	dE3.4	dE3.5
dE1.1	0	15,1528	88,4718	119,783	90,4391	178,624	73,2734	12,7397	71,2644	9,47622	2364,33	4217,12	4362,2	2642,82	3713,61
dE1.2	15,1528	0	30,6744	49,8847	31,8222	97,2428	150,893	54,6729	22,5241	44,0889	2003,55	3729,77	3866,79	2260,13	3258,07
dE1.3	88,4718	30,6744	0	2,74148	0,53352	37,0516	316,738	166,657	6,75002	146,902	1539,86	3087,07	3210,65	1766,48	2659,17
dE1.4	119,783	49,8847	2,74148	0	2,30511	25,1106	370,649	208,503	10,8241	184,144	1422,61	2918,63	3040,37	1639,79	2503,67
dE1.5	90,4391	31,8222	0,53352	2,30511	0	30,011	318,597	169,684	4,35699	147,415	1531,65	3073,63	3198,76	1756,37	2647,03
dE2.1	178,624	97,2428	37,0516	25,1106	30,011	0	441,555	278,95	27,7451	235,644	1336,96	2773,58	2903,12	1537,45	2372,15
dE2.2	73,2734	150,893	316,738	370,649	318,597	441,555	0	26,7741	273,148	34,4093	3241,6	5338,87	5528,12	3562,28	4793,13
dE2.3	12,7397	54,6729	166,657	208,503	169,684	278,95	26,7741	0	141,606	3,8251	2719	4686,52	4839,55	3016,92	4155,55
dE2.4	71,2644	22,5241	6,75002	10,8241	4,35699	27,7451	273,148	141,606	0	116,568	1638,91	3218,98	3351,22	1867,94	2783,2
dE2.5	9,47622	44,0889	146,902	184,144	147,415	235,644	34,4093	3,8251	116,568	0	2623,89	4556,72	4711,75	2913,6	4034,04
dE3.1	2364,33	2003,55	1539,86	1422,61	1531,65	1336,96	3241,6	2719	1638,91	2623,89	0	268,157	303,777	9,36306	152,725
dE3.2	4217,12	3729,77	3087,07	2918,63	3073,63	2773,58	5338,87	4686,52	3218,98	4556,72	268,157	0	3,4712	183,85	17,4569
dE3.3	4362,2	3866,79	3210,65	3040,37	3198,76	2903,12	5528,12	4839,55	3351,22	4711,75	303,777	3,4712	0	216,403	27,3734
dE3.4	2642,82	2260,13	1766,48	1639,78	1756,37	1537,45	3562,28	3016,92	1867,94	2913,6	9,36306	183,85	216,403	0	92,6185
dE3.5	3713,61	3258,07	2659,17	2503,67	2647,03	2372,15	4793,13	4155,55	2783,2	4034,04	152,725	17,4569	27,3734	92,6185	0
dE4.1	4260,13	3770,6	3122,98	2955,02	3111,26	2820,34	5412,93	4731,91	3261,78	4605,67	277,43	2,29262	0,79647	193,975	20,4054
dE4.2	1874,44	1555,24	1151,46	1049,69	1142,78	969,428	2660,98	2192,1	1232,28	2103,36	30,5535	469,605	520,197	67,4525	312,146
dE4.3	1371,99	1100,68	765,126	683,139	758,917	629,438	2057,1	1645,12	834,904	1571,15	134,843	779,175	841,891	207,299	527,041
dE4.4	1622,79	1326,31	955,052	862,882	948,031	798,705	2360,59	1918,84	1032,47	1838,34	69,9324	608,5	664,244	124,381	427,344
dE4.5	1326,68	1059,97	730,728	651,018	725,697	605,458	2002,44	1595,06	802,43	1524,53	149,458	815,773	878,292	226,528	603,497
dE5.1	385,519	248,863	107,552	77,1868	105,13	62,0366	775,441	535,485	129,43	489,685	849,573	2060,24	2166,51	1015,38	1715,35
dE5.2	1427,81	1150,47	808,176	722,08	800,073	654,512	2118,28	1705,99	874,193	1626,27	121,765	739,795	804,83	187,644	539,858
dE5.3	271,51	166,102	70,4798	49,5022	62,1638	15,1283	591,262	396,327	67,4753	346,728	1115,64	2445,68	2568,09	1298,13	2069,59
dE5.4	13,4675	0,11596	33,2966	53,1251	34,2865	99,8823	145,627	51,5966	23,9387	40,9634	2023,1	3756,11	3893,83	2280,78	3282,56
dE5.5	261,813	153,881	52,5733	33,0592	47,8708	21,9869	594,455	388,221	62,5954	346,809	1072,47	2398,68	2513,3	1258,75	2023,08

Figura 1: Tabela que ilustra a matriz de (dis)similaridade dos sinais referentes aos comandos.

Por meio da análise das distâncias euclidianas, é possível perceber que, ao se comparar comandos diferentes, praticamente não existe semelhança entre estes. Enquanto que os mesmos comandos apresentam uma similaridade bem mais significativa quando comparados entre si. Dessa forma, nota-se um pequeno erro associado ao se identificar e diferenciar um comando do outro (fonemas semelhantes), mostrando a validade do método de decisão.

Referências

- [1] E. Thro. The artificial intelligence dictionary, *The Lance. A. Leventhal Microtrend Series*, San Marcos: Microtrend, 1991.
- [2] W. Bussab, E. Miazaki e D.F. Andrade. Introdução à análise de agrupamentos, *Associação Brasileira de Estatística*. 9º Simpósio Nacional de Probabilidade e Estatística, São Paulo, 1990.