

Estimación de nota musical

Aprendizaje Automatizado

Saul Ivan Rivas Vega

A series of horizontal stripes in various colors (beige, blue, red, orange) spanning the width of the slide, located at the bottom.

Descripción y objetivos

Basado en atributos en el dominio de la frecuencia como son los coeficientes por ventana de muestreo de la transformada **constante-Q** estimar la nota musical de guitarra presente en un archivo de audio de una pieza musical monofónica con notas simples.

- Extraer los atributos frecuenciales por ventana del conjunto de datos.
- Entrenar un modelo predictivo para estimar la nota musical presente en una ventana dados sus atributos frecuenciales.
- Realizar un transcript de alguna melodía pre-grabada.

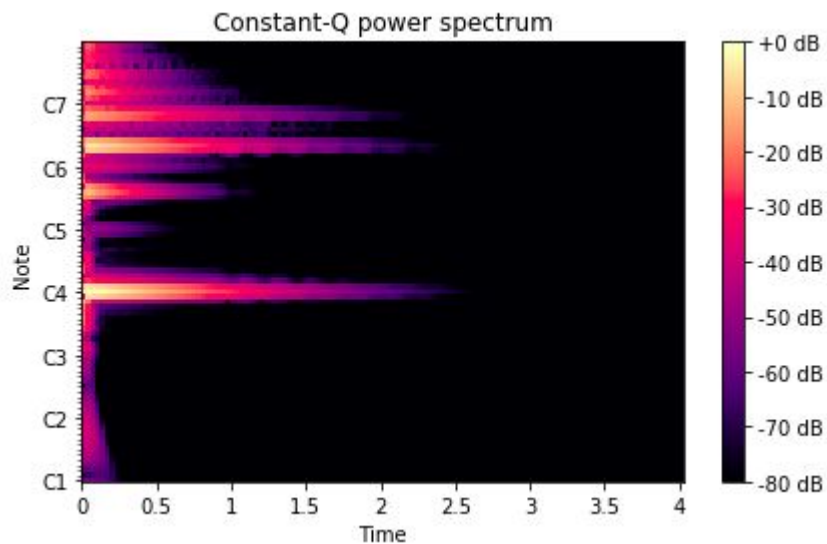
Justificación y Dataset

Poder utilizar las propiedades fundamentales para la estimación del tono y a su vez beneficiarse de los métodos de aprendizaje automatizado para ofrecer una opción para la estimación de notas musicales balanceando la eficiencia en la cantidad de atributos requeridos y la precisión de la estimación.

NSynth contiene 305,979 clips de audio de notas musicales, cada una con un distinto tono, timbre, y envoltura, obtenidos de 1,006 instrumentos grabando clips de monofónicos con una tasa de muestreo de 16kHz de 4 segundos con anotaciones de nota musical en el rango del formato MIDI (21-108) con 5 velocidades (25, 50, 75, 100, 127).



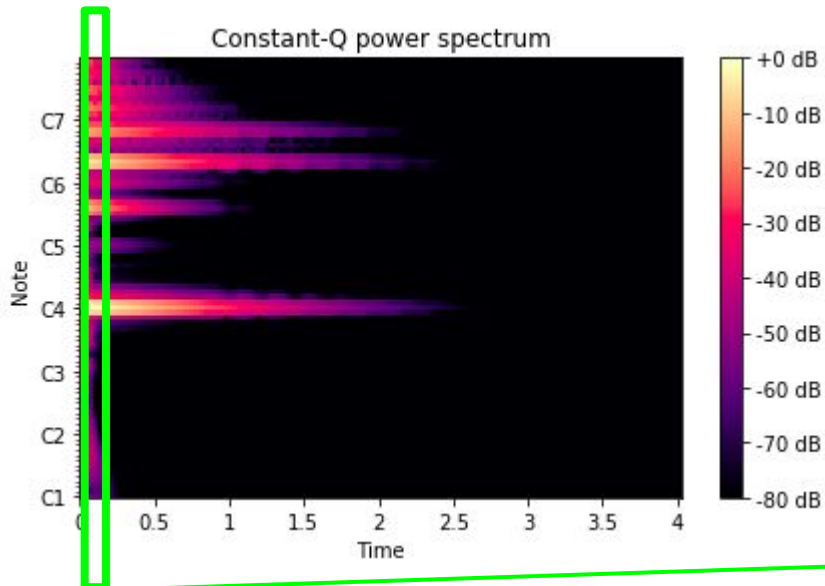
814 Clips de audio de 37 notas desde C1 a C4 provenientes de 22 instrumentos etiquetados como guitarras acústicas. De cada uno se tomaron los primeros 4 coeficientes dando un total de 3256 registros.



	segment_name	note_24	note_25
0	guitar_acoustic_001-042-127_seg_0	0.1548396758272465	0.21162235488528283
1	guitar_acoustic_001-042-127_seg_1	0.14865352420076297	0.20248682732713844
2	guitar_acoustic_001-042-127_seg_2	0.13246161845622467	0.17783167815307035
3	guitar_acoustic_001-042-127_seg_3	0.10867733895496602	0.14073855298161575
4	guitar_acoustic_004-033-127_seg_0	0.2665652028886773	0.43013401153111186
...
3251	guitar_acoustic_026-060-127_seg_3	0.3128918375472502	0.31833682526523954
3252	guitar_acoustic_002-042-127_seg_0	0.45726032931360555	0.4228486327314009
3253	guitar_acoustic_002-042-127_seg_1	0.4387334557786869	0.4036912168580886
3254	guitar_acoustic_002-042-127_seg_2	0.38990515722765556	0.35416741550726255
3255	guitar_acoustic_002-042-127_seg_3	0.3178505825897622	0.2831098900920131

3256 rows × 85 columns

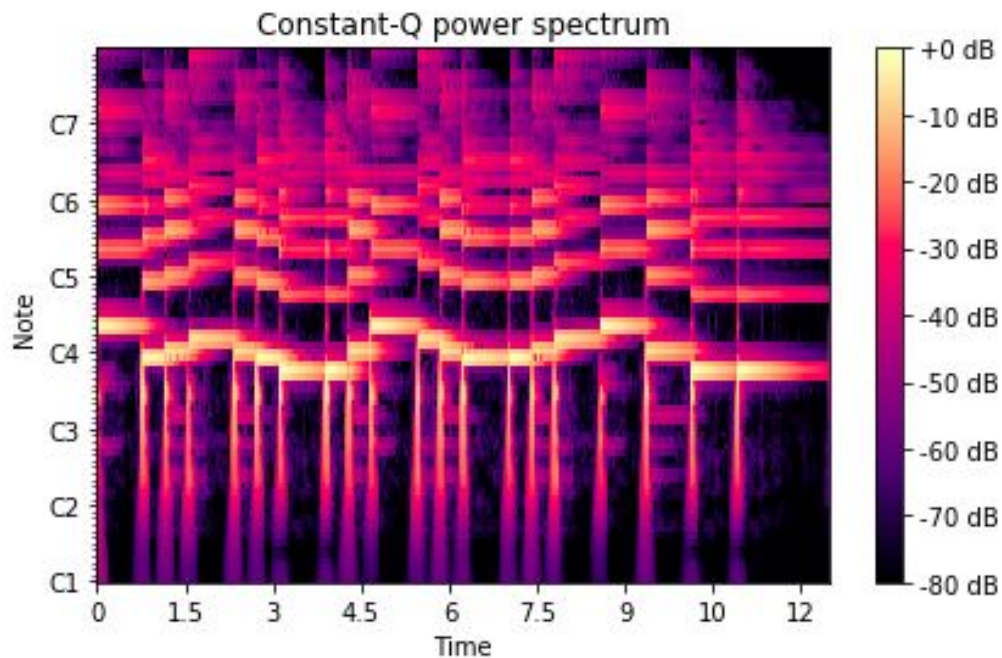
814 Clips de audio de 37 notas desde C1 a C4 provenientes de 22 instrumentos etiquetados como guitarras acústicas. De cada uno se tomaron los primeros 4 coeficientes dando un total de 3256 registros.



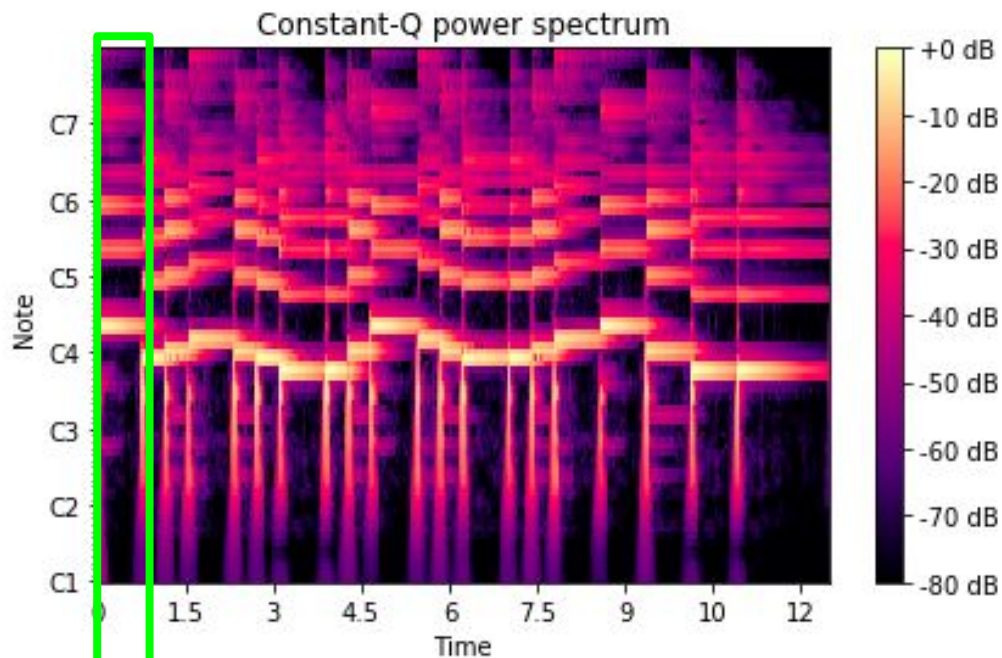
	segment_name	note_24	note_25
0	guitar_acoustic_001-042-127_seg_0	0.1548396758272465	0.21162235488528283
1	guitar_acoustic_001-042-127_seg_1	0.14865352420076297	0.20248682732713844
2	guitar_acoustic_001-042-127_seg_2	0.13246161845622467	0.17783167815307035
3	guitar_acoustic_001-042-127_seg_3	0.10867733895496602	0.14073855298161575
4	guitar_acoustic_004-033-127_seg_0	0.2665652028886773	0.43013401153111186
...
3251	guitar_acoustic_026-060-127_seg_3	0.3128918375472502	0.31833682526523954
3252	guitar_acoustic_002-042-127_seg_0	0.45726032931360555	0.4228486327314009
3253	guitar_acoustic_002-042-127_seg_1	0.4387334557786869	0.4036912168580886
3254	guitar_acoustic_002-042-127_seg_2	0.38990515722765556	0.35416741550726255
3255	guitar_acoustic_002-042-127_seg_3	0.3178505825897622	0.2831098900920131

3256 rows × 85 columns

note_106	note_107	NOTE_CLASS
5.105543487994791e-05	2.11414371213532e-05	42
0.008667981941052722	0.003155229680306545	42
0.0006787049657676877	0.0013428408854527668	42
0.0015128041767926825	0.0002903539822933042	42
7.979831352343817e-07	1.6047887803798125e-07	33
...
0.03701983165047271	0.04718326573072677	60
3.389439542710119e-06	2.0961770858611154e-06	42
0.06745135639540611	0.05517935759730043	42
0.02173963034875114	0.01034854306201443	42
0.00335268698271044	0.0025471661892740916	42



note_106	note_107	NOTE_CLASS
5.105543487994791e-05	2.11414371213532e-05	42
0.008667981941052722	0.003155229680306545	42
0.0006787049657676877	0.0013428408854527668	42
0.0015128041767926825	0.0002903539822933042	42
7.979831352343817e-07	1.6047887803798125e-07	33
...
0.03701983165047271	0.04718326573072677	60
3.389439542710119e-06	2.09617708586111154e-06	42
0.06745135639540611	0.05517935759730043	42
0.02173963034875114	0.01034854306201443	42
0.00335268698271044	0.0025471661892740916	42



X = (0.213124,...,1.62345)

Y = ?

Referencias

- [1] O. Das, J. O. S. Iii, and C. Chafe, "Real-time Pitch Tracking in Audio Signals with the Extended Complex Kalman Filter," p. 7, 2017.
- [2] J. W. Kim, J. Salamon, P. Li, and J. P. Bello, "CREPE: A Convolutional Representation for Pitch Estimation," *arXiv:1802.06182 [cs, eess, stat]*, Feb. 2018, arXiv: 1802.06182. [Online]. Available: <http://arxiv.org/abs/1802.06182>
- [3] M. Mauch and S. Dixon, "PYIN: A fundamental frequency estimator using probabilistic threshold distributions," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 659–663, iSSN: 2379-190X.
- [4] A. de Cheveigné and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1917–1930, Apr. 2002, publisher: Acoustical Society of America. [Online]. Available: <https://asa.scitation.org/doi/10.1121/1.1458024>
- [5] D. Bogdanov, N. Wack, E. Gómez, S. Gulati, P. Herrera, O. Mayor, G. Roma, J. Salamon, J. Zapata, and X. Serra, "ESSENTIA: an Audio Analysis Library for Music Information Retrieval," Nov. 2013.
- [6] J. C. Brown, "Calculation of a constant Q spectral transform," *The Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 425–434, Jan. 1991. [Online]. Available: <http://asa.scitation.org/doi/10.1121/1.400476>
- [7] J. Brown and M. Puckette, "An efficient algorithm for the calculation of a constant Q transform," *Journal of the Acoustical Society of America*, vol. 92, p. 2698, Nov. 1992.
- [8] J. Engel, C. Resnick, A. Roberts, S. Dieleman, D. Eck, K. Simonyan, and M. Norouzi, "Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders," *arXiv:1704.01279 [cs]*, Apr. 2017, arXiv: 1704.01279. [Online]. Available: <http://arxiv.org/abs/1704.01279>
- [9] B. McFee, C. Raffel, D. Liang, D. P. W. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and Music Signal Analysis in Python," Jan. 2015. [Online]. Available: <https://scinapse.io/papers/2191779130>

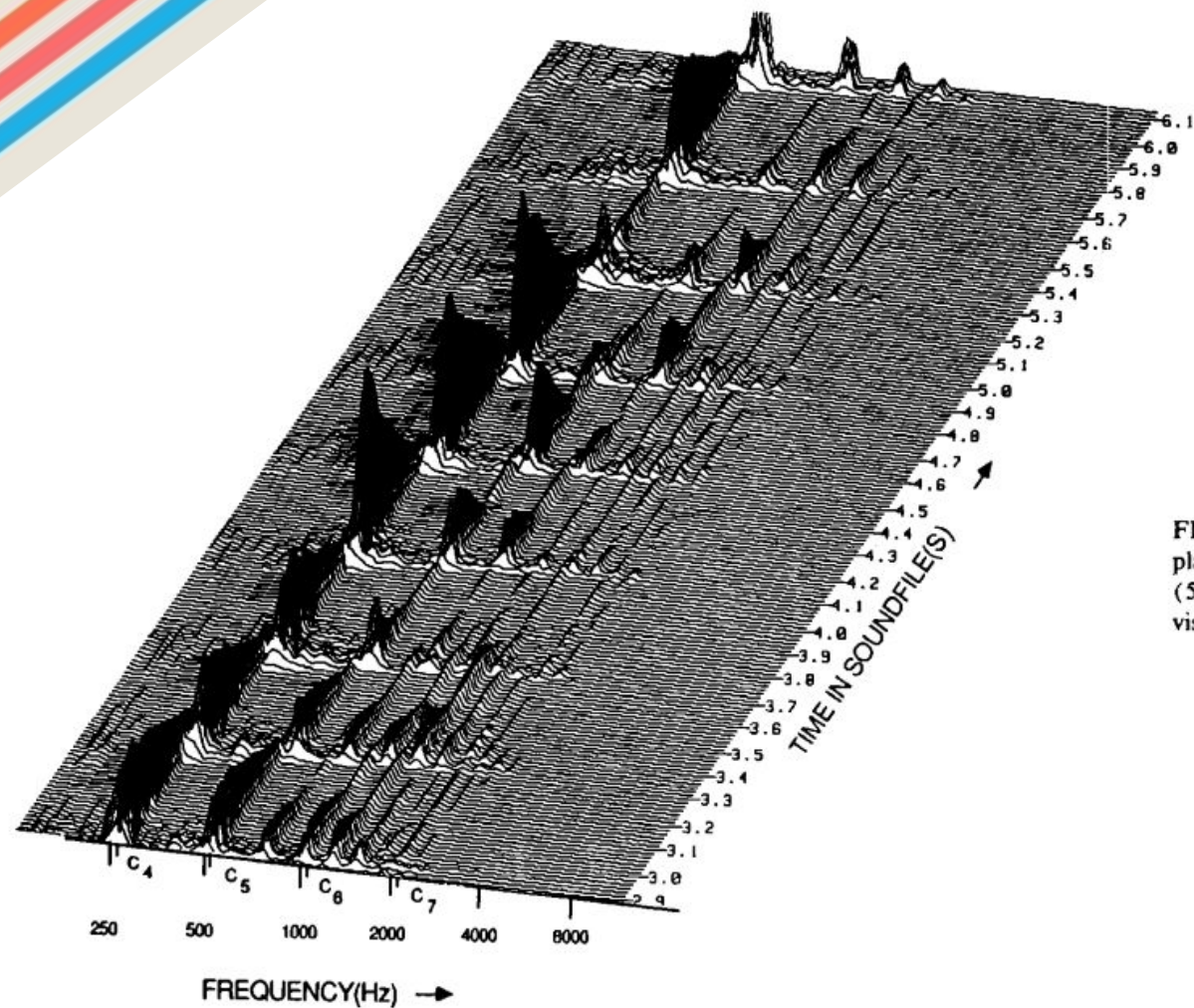


FIG. 10. Constant Q transform of piano playing diatonic scale from C₄ (262 Hz) to C₅ (523 Hz). The attack on D₅ (587 Hz) is also visible. [2]