

The Lakh MIDI Dataset: How it was made, and how to use it

Colin Raffel

C4DM Seminar, November 30, 2016

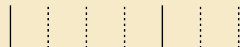
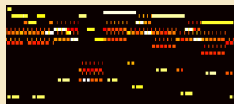


IGERT Integrative Graduate
Education and Research Traineeship



MIDI Files

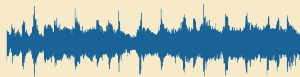
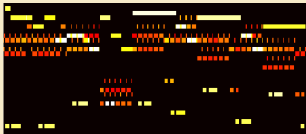
0.0s: F major, 53.2s: D minor, ...



0.0s: 110 bpm, 22.2s: 103 bpm, ...

The Goal

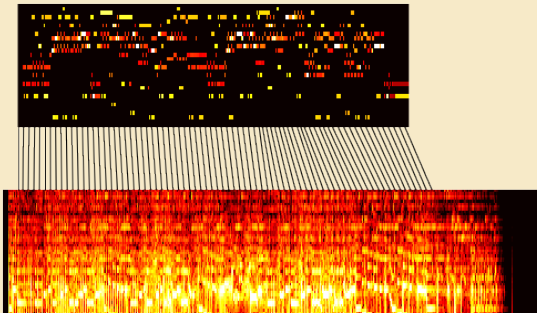
```
artist: 'Tori Amos'  
release: 'LIVE AT MONTREUX'  
title: 'Smells Like Teen Spirit'  
id: 'TRKUYPW128F92E1FC0'  
duration: 216.4502  
sample_rate: 22050  
audio_md5: '8'  
7digitalid: 5764727  
year: 1992
```



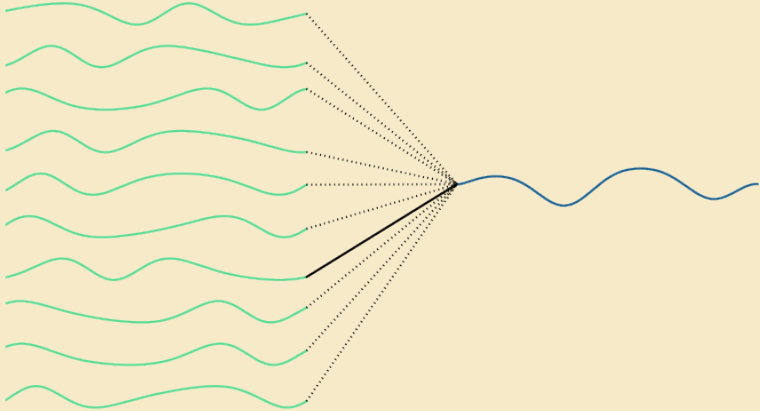
Matching and Aligning



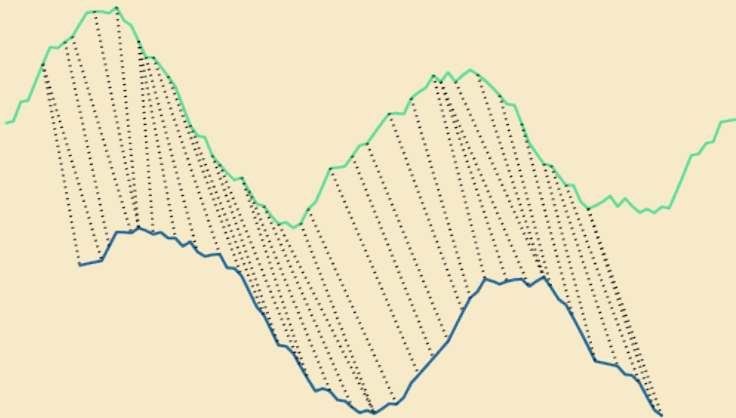
Beatles/hello, goodbye.mid ↔ The Beatles/Hello.mp3



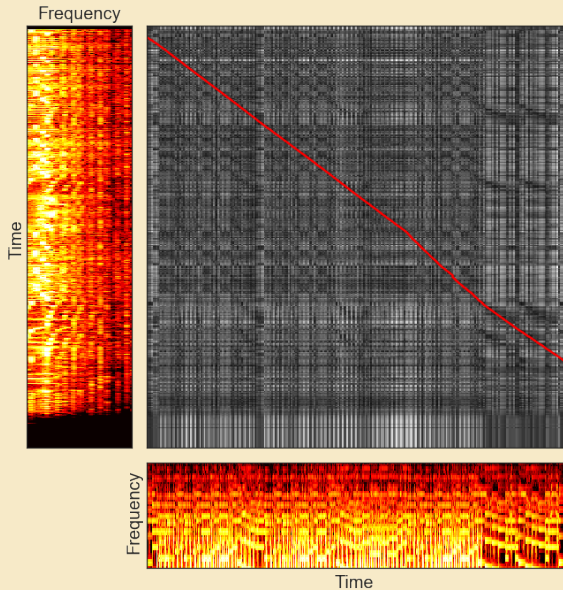
Sequence Matching



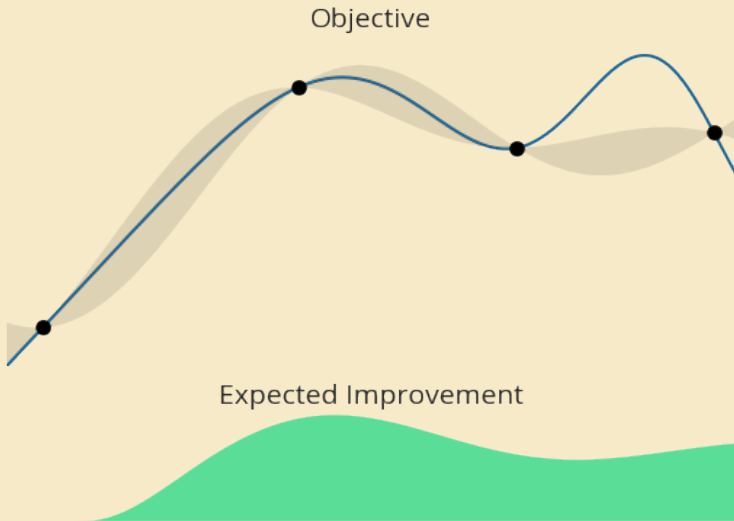
Dynamic Time Warping



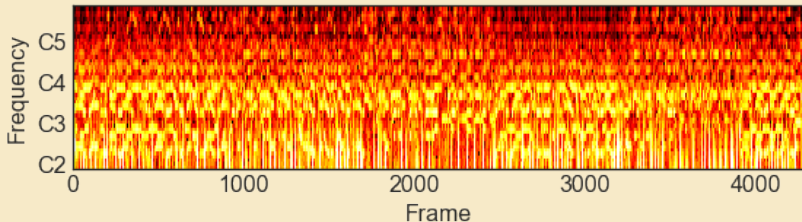
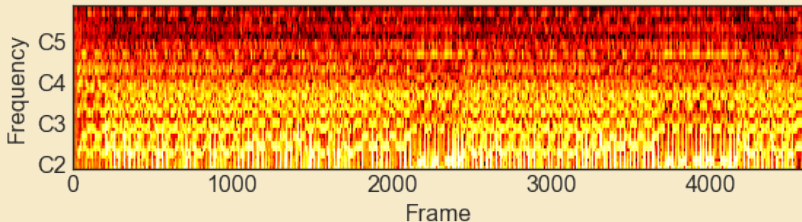
Aligning MIDI Files with DTW



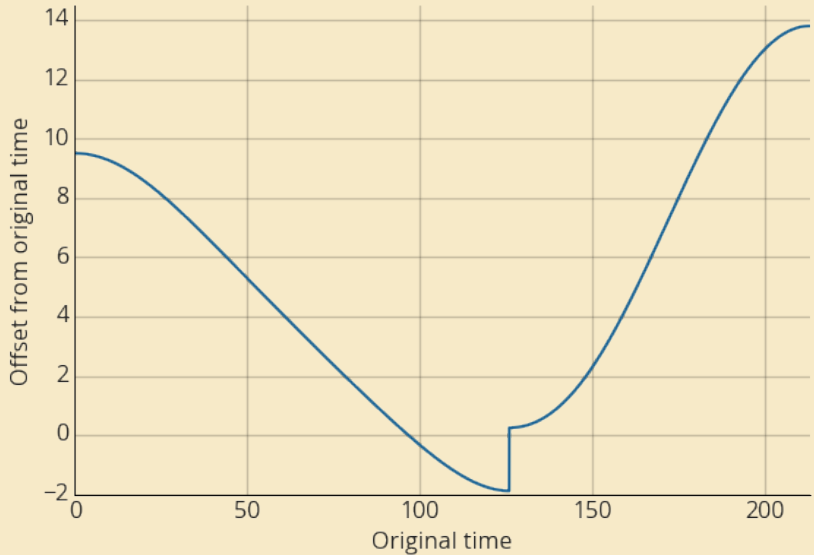
Bayesian Optimization



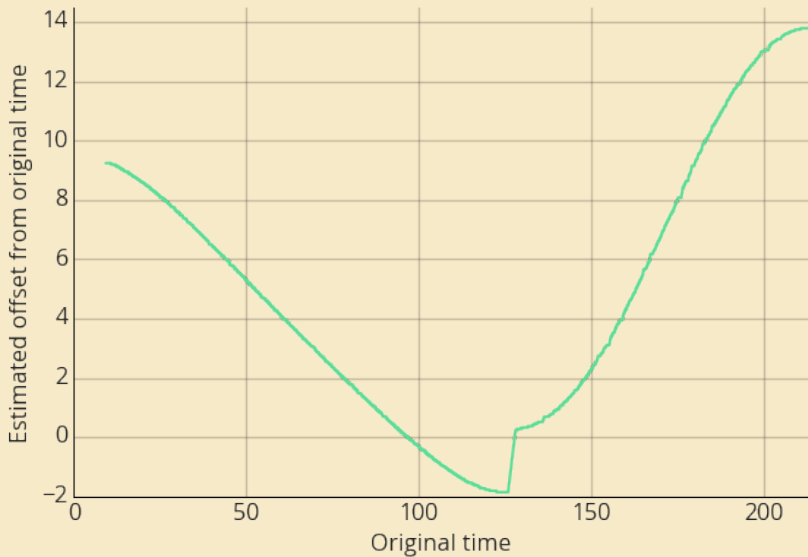
Synthetic Alignment Data



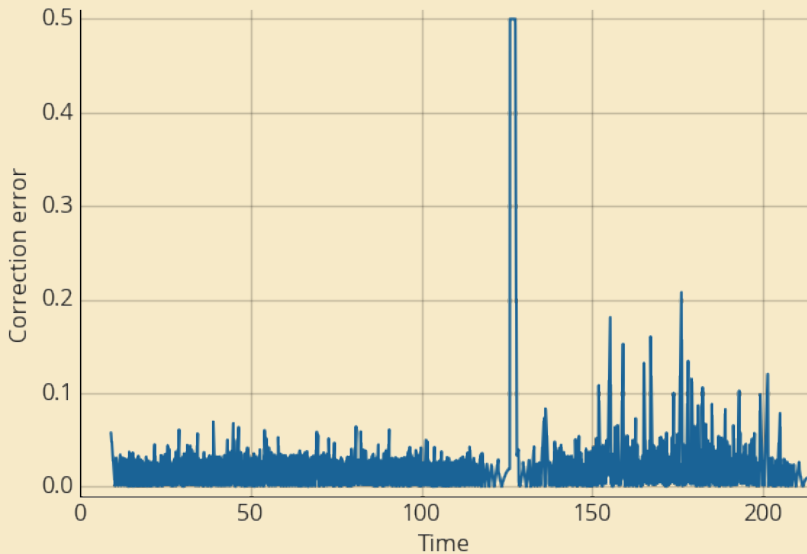
Artificial Time Warping



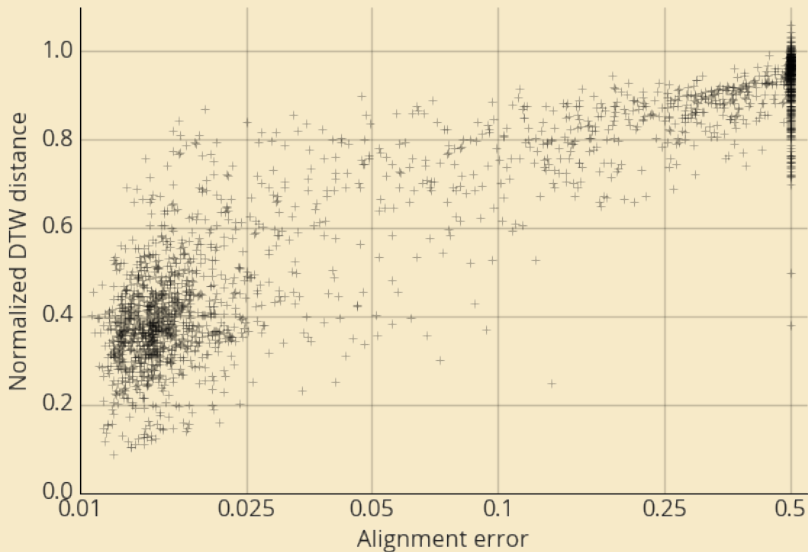
Correcting Time Warping



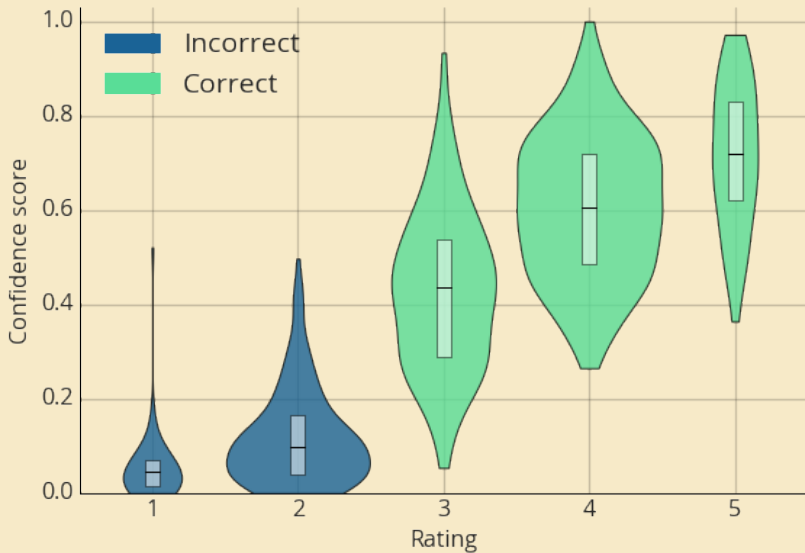
Measuring Error



Score Normalization Search



Real-World Test



DTW is Too Slow

(.247 seconds)

DTW is Too Slow

(.247 seconds)(178,561 MIDI files)

DTW is Too Slow

(.247 seconds)(178,561 MIDI files)(994,960 audio files)

DTW is Too Slow

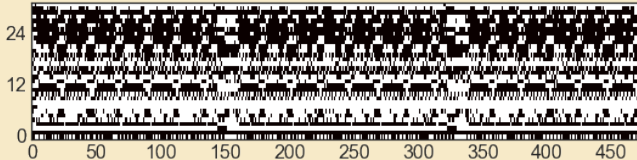
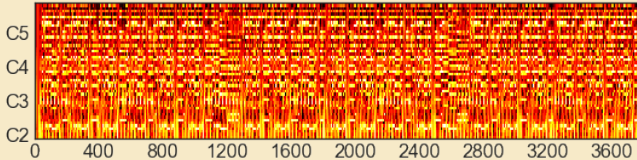
$$\frac{(.247 \text{ seconds})(178,561 \text{ MIDI files})(994,960 \text{ audio files})}{(60 \text{ seconds})(60 \text{ minutes})(24 \text{ hours})(365 \text{ days})}$$

DTW is Too Slow

$$\frac{(.247 \text{ seconds})(178,561 \text{ MIDI files})(994,960 \text{ audio files})}{(60 \text{ seconds})(60 \text{ minutes})(24 \text{ hours})(365 \text{ days})}$$

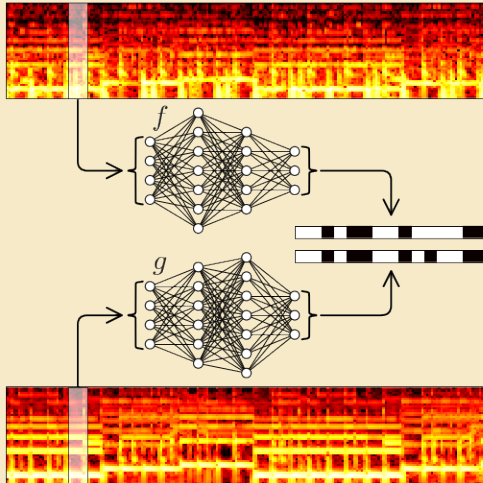
$\approx 1,391$ years

Downsampled Hash Sequences

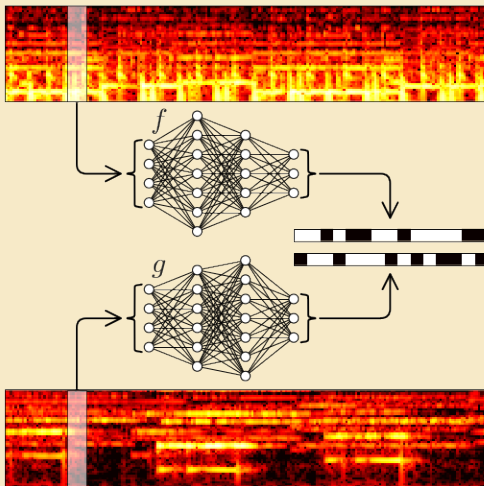


A diagram illustrating the POPCNT (Population Count) operation. It shows three horizontal bars representing binary sequences. The top bar has black segments at positions 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 23, 25, 26, 28, 29, 31, 32, 34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52, 53, 55, 56, 58, 59, 61, 62, 64, 65, 67, 68, 70, 71, 73, 74, 76, 77, 79, 80, 82, 83, 85, 86, 88, 89, 91, 92, 94, 95, 97, 98, 100, 101, 103, 104, 106, 107, 109, 110, 112, 113, 115, 116, 118, 119, 121, 122, 124, 125, 127, 128, 130, 131, 133, 134, 136, 137, 139, 140, 142, 143, 145, 146, 148, 149, 151, 152, 154, 155, 157, 158, 160, 161, 163, 164, 166, 167, 169, 170, 172, 173, 175, 176, 178, 179, 181, 182, 184, 185, 187, 188, 190, 191, 193, 194, 196, 197, 199, 200, 202, 203, 205, 206, 208, 209, 211, 212, 214, 215, 217, 218, 220, 221, 223, 224, 226, 227, 229, 230, 232, 233, 235, 236, 238, 239, 241, 242, 244, 245, 247, 248, 250, 251, 253, 254, 256, 257, 259, 260, 262, 263, 265, 266, 268, 269, 271, 272, 274, 275, 277, 278, 280, 281, 283, 284, 286, 287, 289, 290, 292, 293, 295, 296, 298, 299, 301, 302, 304, 305, 307, 308, 310, 311, 313, 314, 316, 317, 319, 320, 322, 323, 325, 326, 328, 329, 331, 332, 334, 335, 337, 338, 340, 341, 343, 344, 346, 347, 349, 350, 352, 353, 355, 356, 358, 359, 361, 362, 364, 365, 367, 368, 370, 371, 373, 374, 376, 377, 379, 380, 382, 383, 385, 386, 388, 389, 391, 392, 394, 395, 397, 398, 400, 401, 403, 404, 406, 407, 409, 410, 412, 413, 415, 416, 418, 419, 421, 422, 424, 425, 427, 428, 430, 431, 433, 434, 436, 437, 439, 440, 442, 443, 445, 446, 448, 449, 451, 452, 454, 455, 457, 458, 460, 461, 463, 464, 466, 467, 469, 470, 472, 473, 475, 476, 478, 479, 481, 482, 484, 485, 487, 488, 490, 491, 493, 494, 496, 497, 499, 500, 502, 503, 505, 506, 508, 509, 511, 512, 514, 515, 517, 518, 520, 521, 523, 524, 526, 527, 529, 530, 532, 533, 535, 536, 538, 539, 541, 542, 544, 545, 547, 548, 550, 551, 553, 554, 556, 557, 559, 560, 562, 563, 565, 566, 568, 569, 571, 572, 574, 575, 577, 578, 580, 581, 583, 584, 586, 587, 589, 590, 592, 593, 595, 596, 598, 599, 601, 602, 604, 605, 607, 608, 610, 611, 613, 614, 616, 617, 619, 620, 622, 623, 625, 626, 628, 629, 631, 632, 634, 635, 637, 638, 640, 641, 643, 644, 646, 647, 649, 650, 652, 653, 655, 656, 658, 659, 661, 662, 664, 665, 667, 668, 670, 671, 673, 674, 676, 677, 679, 680, 682, 683, 685, 686, 688, 689, 691, 692, 694, 695, 697, 698, 700, 701, 703, 704, 706, 707, 709, 710, 712, 713, 715, 716, 718, 719, 721, 722, 724, 725, 727, 728, 730, 731, 733, 734, 736, 737, 739, 740, 742, 743, 745, 746, 748, 749, 751, 752, 754, 755, 757, 758, 760, 761, 763, 764, 766, 767, 769, 770, 772, 773, 775, 776, 778, 779, 781, 782, 784, 785, 787, 788, 790, 791, 793, 794, 796, 797, 799, 800, 802, 803, 805, 806, 808, 809, 811, 812, 814, 815, 817, 818, 820, 821, 823, 824, 826, 827, 829, 830, 832, 833, 835, 836, 838, 839, 841, 842, 844, 845, 847, 848, 850, 851, 853, 854, 856, 857, 859, 860, 862, 863, 865, 866, 868, 869, 871, 872, 874, 875, 877, 878, 880, 881, 883, 884, 886, 887, 889, 890, 892, 893, 895, 896, 898, 899, 901, 902, 904, 905, 907, 908, 910, 911, 913, 914, 916, 917, 919, 920, 922, 923, 925, 926, 928, 929, 931, 932, 934, 935, 937, 938, 940, 941, 943, 944, 946, 947, 949, 950, 952, 953, 955, 956, 958, 959, 961, 962, 964, 965, 967, 968, 970, 971, 973, 974, 976, 977, 979, 980, 982, 983, 985, 986, 988, 989, 991, 992, 994, 995, 997, 998, 1000, 1002, 1003, 1005, 1006, 1008, 1009, 1011, 1012, 1014, 1015, 1017, 1018, 1020, 1021, 1023, 1024, 1026, 1027, 1029, 1030, 1032, 1033, 1035, 1036, 1038, 1039, 1041, 1042, 1044, 1045, 1047, 1048, 1050, 1051, 1053, 1054, 1056, 1057, 1059, 1060, 1062, 1063, 1065, 1066, 1068, 1069, 1071, 1072, 1074, 1075, 1077, 1078, 1080, 1081, 1083, 1084, 1086, 1087, 1089, 1090, 1092, 1093, 1095, 1096, 1098, 1099, 1101, 1102, 1104, 1105, 1107, 1108, 1110, 1111, 1113, 1114, 1116, 1117, 1119, 1120, 1122, 1123, 1125, 1126, 1128, 1129, 1131, 1132, 1134, 1135, 1137, 1138, 1140, 1141, 1143, 1144, 1146, 1147, 1149, 1150, 1152, 1153, 1155, 1156, 1158, 1159, 1161, 1162, 1164, 1165, 1167, 1168, 1170, 1171, 1173, 1174, 1176, 1177, 1179, 1180, 1182, 1183, 1185, 1186, 1188, 1189, 1191, 1192, 1194, 1195, 1197, 1198, 1200, 1201, 1203, 1204, 1206, 1207, 1209, 1210, 1212, 1213, 1215, 1216, 1218, 1219, 1221, 1222, 1224, 1225, 1227, 1228, 1230, 1231, 1233, 1234, 1236, 1237, 1239, 1240, 1242, 1243, 1245, 1246, 1248, 1249, 1251, 1252, 1254, 1255, 1257, 1258, 1260, 1261, 1263, 1264, 1266, 1267, 1269, 1270, 1272, 1273, 1275, 1276, 1278, 1279, 1281, 1282, 1284, 1285, 1287, 1288, 1290, 1291, 1293, 1294, 1296, 1297, 1299, 1300, 1302, 1303, 1305, 1306, 1308, 1309, 1311, 1312, 1314, 1315, 1317, 1318, 1320, 1321, 1323, 1324, 1326, 1327, 1329, 1330, 1332, 1333, 1335, 1336, 1338, 1339, 1341, 1342, 1344, 1345, 1347, 1348, 1350, 1351, 1353, 1354, 1356, 1357, 1359, 1360, 1362, 1363, 1365, 1366, 1368, 1369, 1371, 1372, 1374, 1375, 1377, 1378, 1380, 1381, 1383, 1384, 1386, 1387, 1389, 1390, 1392, 1393, 1395, 1396, 1398, 1399, 1401, 1402, 1404, 1405, 1407, 1408, 1410, 1411, 1413, 1414, 1416, 1417, 1419, 1420, 1422, 1423, 1425, 1426, 1428, 1429, 1431, 1432, 1434, 1435, 1437, 1438, 1440, 1441, 1443, 1444, 1446, 1447, 1449, 1450, 1452, 1453, 1455, 1456, 1458, 1459, 1461, 1462, 1464, 1465, 1467, 1468, 1470, 1471, 1473, 1474, 1476, 1477, 1479, 1480, 1482, 1483, 1485, 1486, 1488, 1489, 1491, 1492, 1494, 1495, 1497, 1498, 1500, 1501, 1503, 1504, 1506, 1507, 1509, 1510, 1512, 1513, 1515, 1516, 1518, 1519, 1521, 1522, 1524, 1525, 1527, 1528, 1530, 1531, 1533, 1534, 1536, 1537, 1539, 1540, 1542, 1543, 1545, 1546, 1548, 1549, 1551, 1552, 1554, 1555, 1557, 1558, 1560, 1561, 1563, 1564, 1566, 1567, 1569, 1570, 1572, 1573, 1575, 1576, 1578, 1579, 1581, 1582, 1584, 1585, 1587, 1588, 1590, 1591, 1593, 1594, 1596, 1597, 1599, 1600, 1602, 1603, 1605, 1606, 1608, 1609, 1611, 1612, 1614, 1615, 1617, 1618, 1620, 1621, 1623, 1624, 1626, 1627, 1629, 1630, 1632, 1633, 1635, 1636, 1638, 1639, 1641, 1642, 1644, 1645, 1647, 1648, 1650, 1651, 1653, 1654, 1656, 1657, 1659, 1660, 1662, 1663, 1665, 1666, 1668, 1669, 1671, 1672, 1674, 1675, 1677, 1678, 1680, 1681, 1683, 1684, 1686, 1687, 1689, 1690, 1692, 1693, 1695, 1696, 1698, 1699, 1701, 1702, 1704, 1705, 1707, 1708, 1710, 1711, 1713, 1714, 1716, 1717, 1719, 1720, 1722, 1723, 1725, 1726, 1728, 1729, 1731, 1732, 1734, 1735, 1737, 1738, 1740, 1741, 1743, 1744, 1746, 1747, 1749, 1750, 1752, 1753, 1755, 1756, 1758, 1759, 1761, 1762, 1764, 1765, 1767, 1768, 1770, 1771, 1773, 1774, 1776, 1777, 1779, 1780, 1782, 1783, 1785, 1786, 1788, 1789, 1791, 1792, 1794, 1795, 1797, 1798, 1800, 1801, 1803, 1804, 1806, 1807, 1809, 1810, 1812, 1813, 1815, 1816, 1818, 1819, 1821, 1822, 1824, 1825, 1827, 1828, 1830, 1831, 1833, 1834, 1836, 1837, 1839, 1840, 1842, 1843, 1845, 1846, 1848, 1849, 1851, 1852, 1854, 1855, 1857, 1858, 1860, 1861, 1863, 1864, 1866, 1867, 1869, 1870, 1872, 1873, 1875, 1876, 1878, 1879, 1881, 1882, 1884, 1885, 1887, 1888, 1890, 1891, 1893, 1894, 1896, 1897, 1899, 1900, 1902, 1903, 1905, 1906, 1908, 1909, 1911, 1912, 1914, 1915, 1917, 1918, 1920, 1921, 1923, 1924, 1926, 1927, 1929, 1930, 1932, 1933, 1935, 1936, 1938, 1939, 1941, 1942, 1944, 1945, 1947, 1948, 1950, 1951, 1953, 1954, 1956, 1957, 1959, 1960, 1962, 1963, 1965, 1966, 1968, 1969, 1971, 1972, 1974, 1975, 1977, 1978, 1980, 1981, 1983, 1984, 1986, 1987, 1989, 1990, 1992, 1993, 1995, 1996, 1998, 1999, 2000, 2002, 2003, 2005, 2006, 2008, 2009, 2011, 2012, 2014, 2015, 2017, 2018, 2020, 2021, 2023, 2024, 2026, 2027, 2029, 2030, 2032, 2033, 2035, 2036, 2038, 2039, 2041, 2042, 2044, 2045, 2047, 2048, 2050, 2051, 2053, 2054, 2056, 2057, 2059, 2060, 2062, 2063, 2065, 2066, 2068, 2069, 2071, 2072, 2074, 2075, 2077, 2078, 2080, 2081, 2083, 2084, 2086, 2087, 2089, 2090, 2092, 2093, 2095, 2096, 2098, 2099, 2100, 2102, 2103, 2105, 2106, 2108, 2109, 2111, 2112, 2114, 2115, 2117, 2118, 2120, 2121, 2123, 2124, 2126, 2127, 2129, 2130, 2132, 2133, 2135, 2136, 2138, 2139, 2141, 2142, 2144, 2145, 2147, 2148, 2150, 2151, 2153, 2154, 2156, 2157, 2159, 2160, 2162, 2163, 2165, 2166, 2168, 2169, 2171, 2172, 2174, 2175, 2177, 2178, 2180, 2181, 2183, 2184, 2186, 2187, 2189, 2190, 2192, 2193, 2195, 2196, 2198, 2199, 2200, 2202, 2203, 2205, 2206, 2208, 2209, 2211, 2212, 2214, 2215, 2217, 2218, 2220, 2221, 2223, 2224, 2226, 2227, 2229, 2230, 2232, 2233, 2235, 2236, 2238, 2239, 2241, 2242, 2244, 2245, 2247, 2248, 2250, 2251, 2253, 2254, 2256, 2257, 2259, 2260, 2262, 2263, 2265, 2266, 2268, 2269, 2271, 2272, 2274, 2275, 2277, 2278, 2280, 2281, 2283, 2284, 2286, 2287, 2289, 2290, 2292, 2293, 2295, 2296, 2298, 2299, 2300, 2302, 2303, 2305, 2306, 2308, 2309, 2311, 2312, 2314, 2315, 2317, 2318, 2320, 2321, 2323, 2324, 2326, 2327, 2329, 2330, 2332, 2333, 2335, 2336, 2338, 2339, 2341, 2342, 2344, 2345, 2347, 2348, 2350, 2351, 2353, 2354, 2356, 2357, 2359, 2360, 2362, 2363, 2365, 2366, 2368, 2369, 2371, 2372, 2374, 2375, 2377, 2378, 2380, 2381, 2383, 2384, 2386, 2387, 2389, 2390, 2392, 2393, 2395, 2396, 2398, 2399, 2400, 2402, 2403, 2405, 2406, 2408, 2409, 2411, 2412, 2414, 2415, 2417, 2418, 2420, 2421, 2423, 2424, 2426, 2427, 2429, 2430, 2432, 2433, 2435, 2436, 2438, 2439, 2441, 2442, 2444, 2445, 2447, 2448, 2450, 2451, 2453, 2454, 2456, 2457, 2459, 2460, 2462, 2463, 2465, 2466, 2468, 2469, 2471, 2472, 2474, 2475, 2477, 2478, 2480, 2481, 2483, 2484, 2486, 2487, 2489, 2490, 2492, 2493, 2495, 2496, 2498, 2499, 2500, 2502, 2503, 2505, 2506, 2508, 2509, 2511, 2512, 2514, 2515, 2517, 2518, 2520, 2521, 2523, 2524, 2526, 2527, 2529, 2530, 2532, 2533, 2535, 2536, 2538, 2539, 2541, 2542, 2544, 2545, 2547, 2548, 2550, 2551, 2553, 2554, 2556, 2557, 2559, 2560, 2562, 2563, 2565, 2566, 2568, 2569, 2571, 2572, 2574, 2575, 2577, 2578, 2580, 2581, 2583, 2584, 2586, 2587, 2589, 2590, 2592, 2593, 2595, 2596, 2598, 2599, 2600, 2602, 2603, 2605, 2606, 2608, 2609, 2611, 2612, 2614, 2615, 2617, 2618, 2620, 2621, 2623, 2624, 2626, 2627, 2629, 2630, 2632, 2633, 2635, 2636, 2638, 2639, 2641, 2642, 2644, 2645, 2647, 2648, 2650, 2651, 2653, 2654, 2656, 2657, 2659, 2660, 2662, 2663, 2665, 2666, 2668, 2669, 2671, 2672, 2674, 2675, 2677, 2678, 2680, 2681, 2683, 2684, 2686, 2687, 2689, 2690, 2692, 2693, 2695, 2696, 2698, 2699, 2700, 2702, 2703, 2705, 2706, 2708, 2709, 2711, 2712, 2714, 2715, 2717, 2718, 2720, 2721, 2723, 2724, 2726, 2727, 2729, 2730, 2732, 2733, 2735, 2736, 2738, 2739, 2741, 2742, 2744, 2745, 2747, 2748, 2750, 2751, 2753, 2754, 2756, 2757, 2759, 2760, 2762, 2763, 2765, 2766, 2768, 2769, 2771, 2772, 2774, 2775, 2777, 2778, 2780, 2781, 2783, 2784, 2786, 2787, 2789, 2790, 2792, 2793, 2795, 2796, 2798, 2799, 2800, 2802, 2803, 2805, 2806, 2808, 2809, 2811, 2812, 2814, 2815, 2817, 2818, 2820, 2821, 2823, 2824, 2826, 2827, 2829, 2830, 2832, 2833, 2835, 2836, 2838, 2839, 2841, 2842, 2844, 2845, 2847, 2848, 2850, 2851, 2853, 2854, 2856, 2857, 2859, 2860, 2862, 2863, 2865, 2866, 2868, 2869, 2871, 2872, 2874, 2875, 2877, 2878, 2880, 2881, 2883, 2884, 2886, 2887, 2889, 2890, 2892, 2893, 2895, 2896, 2898, 2899, 2900, 2902, 2903, 2905, 2906, 2908, 2909, 2911, 2912, 2914, 2915, 2917, 2918, 2920, 2921, 2923, 2924, 2926, 2927, 2929, 2930, 2932, 2933, 2935, 2936, 2938, 2939, 2941, 2942, 2944, 2945, 2947, 2948, 2950, 2951, 2953, 2954, 2956, 2957, 2959, 2960, 2962, 2963, 2965, 2966, 2968, 2969, 2971, 2972, 2974, 2975, 2977, 2978, 2980, 2981, 2983, 2984, 2986, 2987, 2989, 2990, 2992, 2993, 2995, 2996, 2998, 2999, 3000, 3002, 3003, 3005, 3006, 3008, 3009, 3011, 3012, 3014, 3015, 3017, 3018, 3020, 3021, 3023, 3024, 3026, 3027, 3029, 3030, 3032, 3033, 3035, 3036, 3038, 3039, 3041, 3042, 3044, 3045, 3047, 3048, 3050, 3051, 3053, 3054, 3056, 3057, 3059, 3060, 3062, 3063, 3065, 3066, 3068, 3069, 3071, 3072, 3074, 3075, 3077, 3078, 3080, 3081, 3083, 3084, 3086, 3087, 3089, 3090, 3092, 3093, 3095, 3096, 3098, 3099, 3100, 3102, 3103, 3105, 3106, 3108, 3109, 3111, 3112, 3114, 3115, 3117, 3118, 3120, 3121, 3123, 3124, 3126, 3127, 3129, 3130, 3132, 3133, 3135, 3136, 3138, 3139, 3141, 3142, 3144, 3145, 3147, 3148, 3150, 3151, 3153, 3154, 3156, 3157,

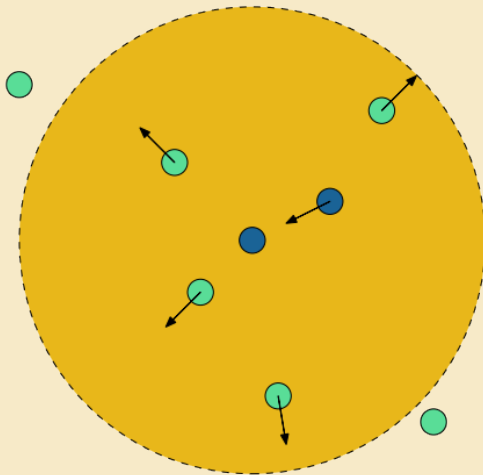
Similarity-Preserving Hashing



Similarity-Preserving Hashing

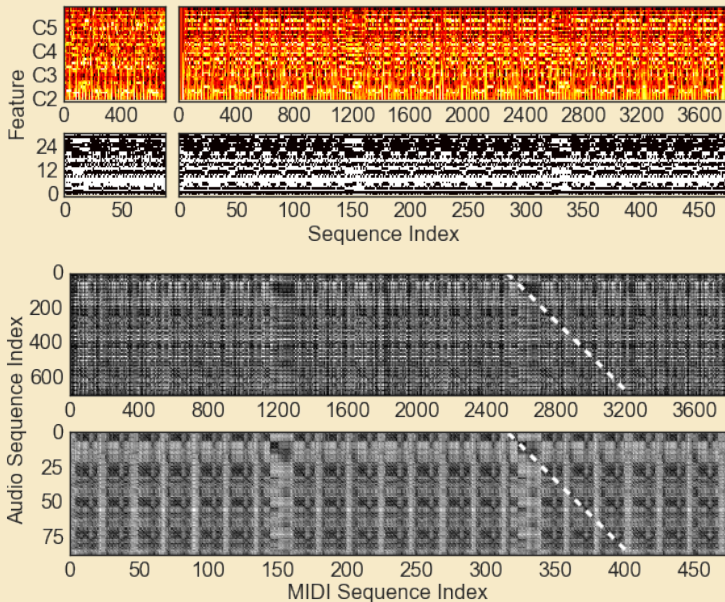


Loss Function

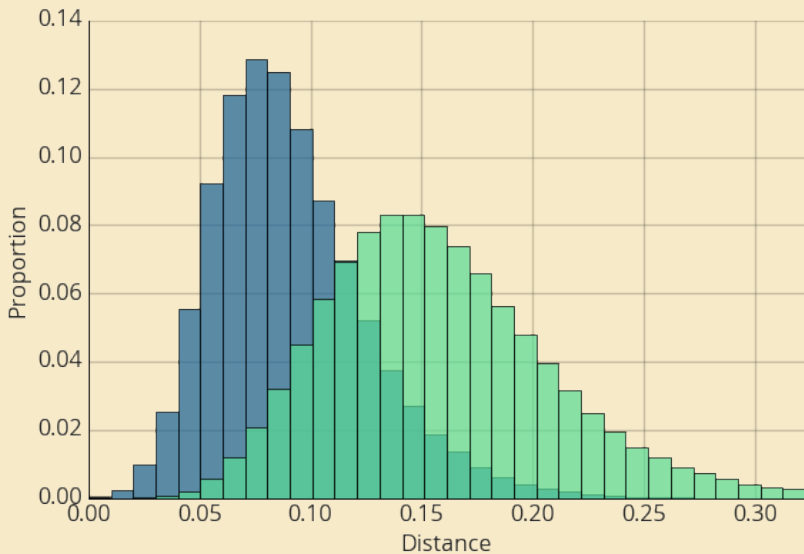


$$\mathcal{L} = \frac{1}{|\mathcal{P}|} \sum_{(x,y) \in \mathcal{P}} \|f(x) - g(y)\|_2^2 + \frac{\alpha}{|\mathcal{N}|} \sum_{(x,y) \in \mathcal{N}} \max(0, m - \|f(x) - g(y)\|_2)^2$$

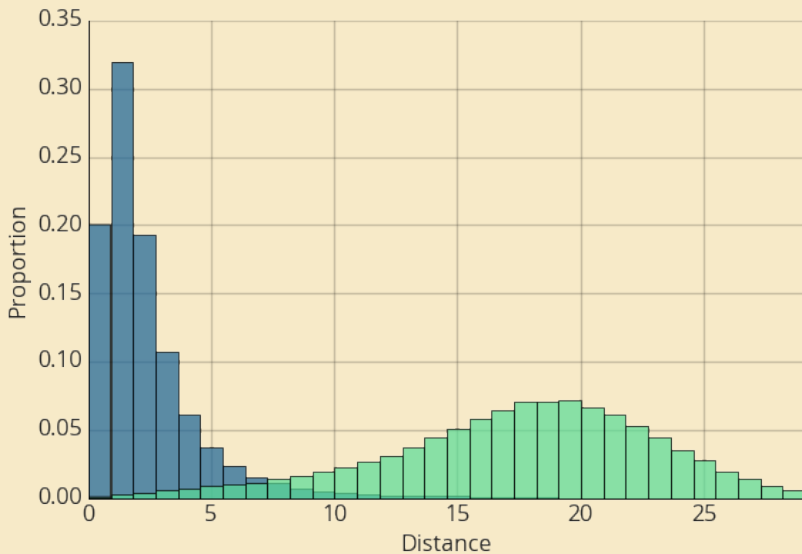
Example Output



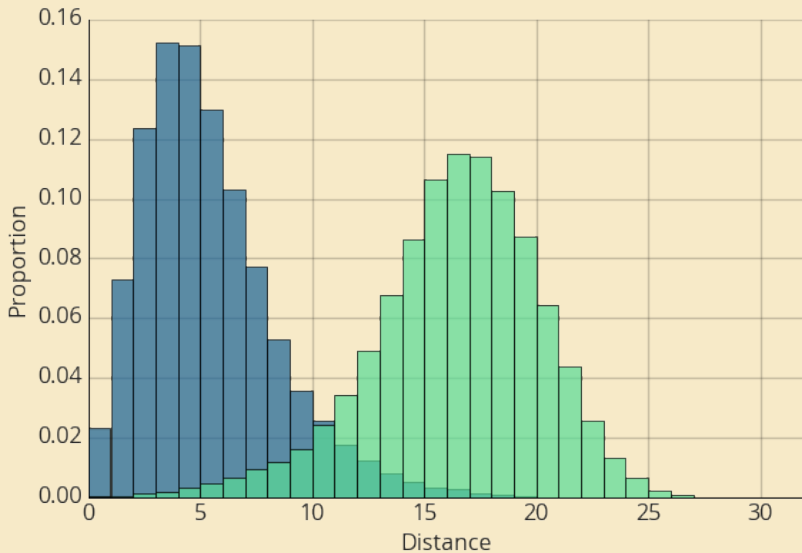
Raw Distance Distributions



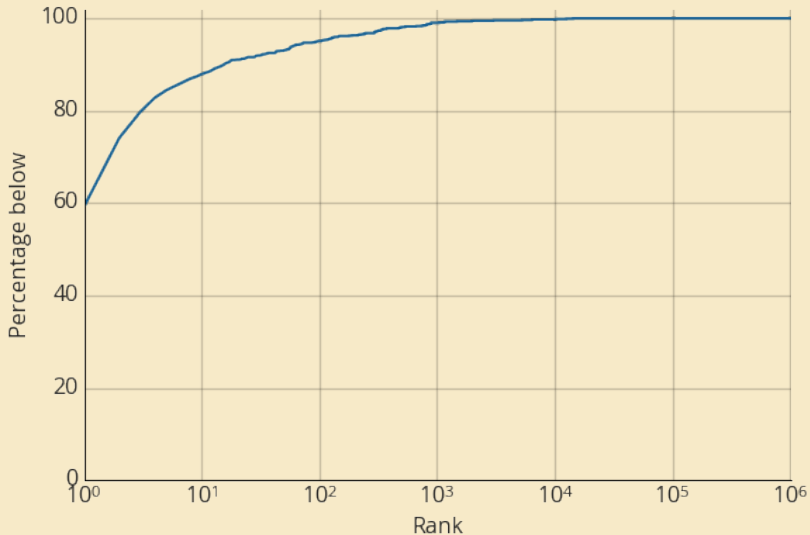
Output Distance Distributions



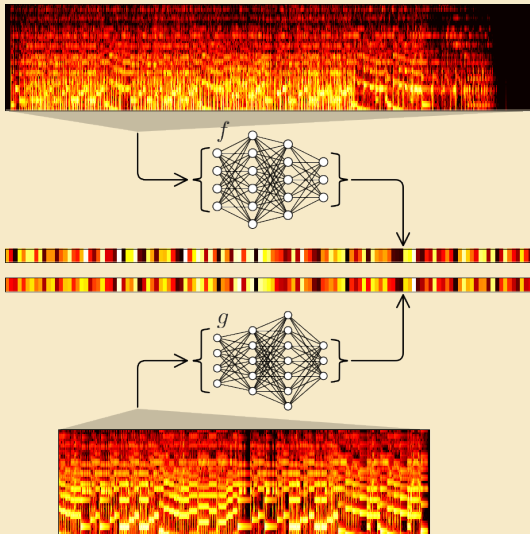
Hash Distance Distributions



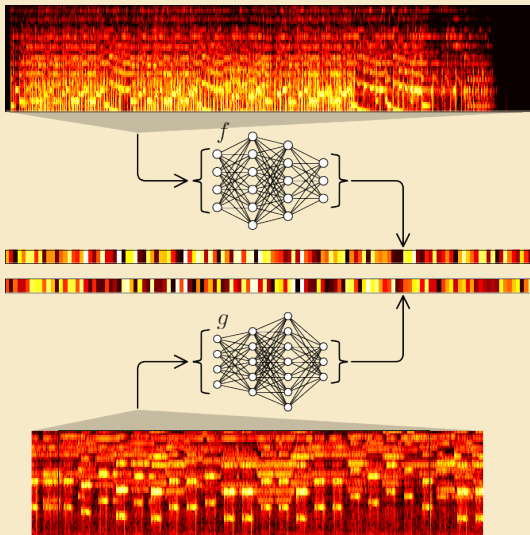
Match Ranks



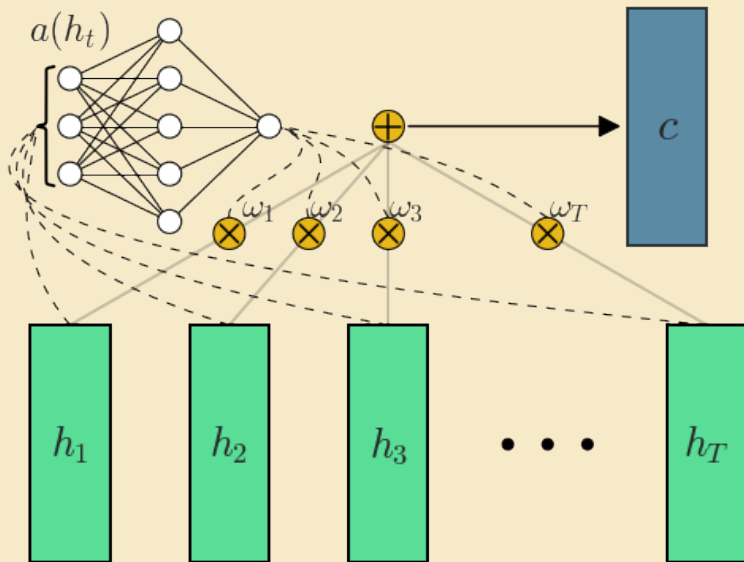
Pairwise Sequence Embedding



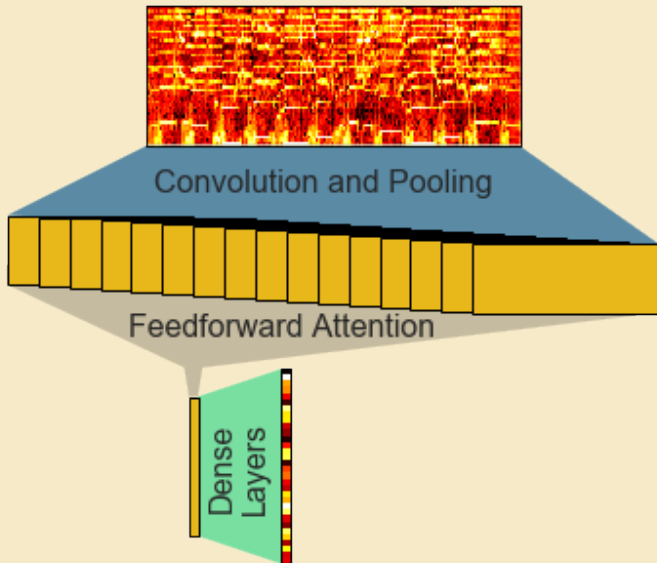
Pairwise Sequence Embedding



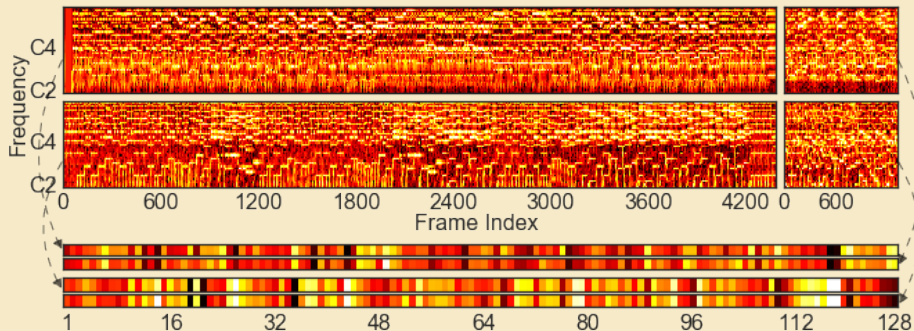
Feed-Forward Attention



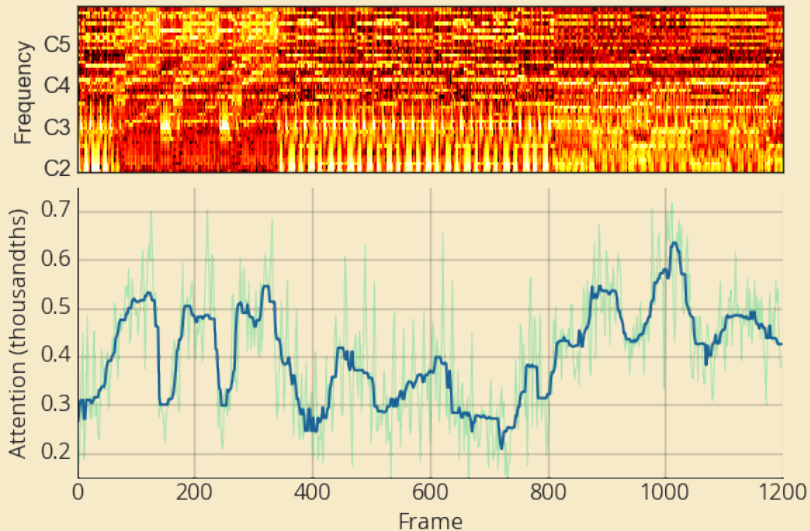
Embedding Network



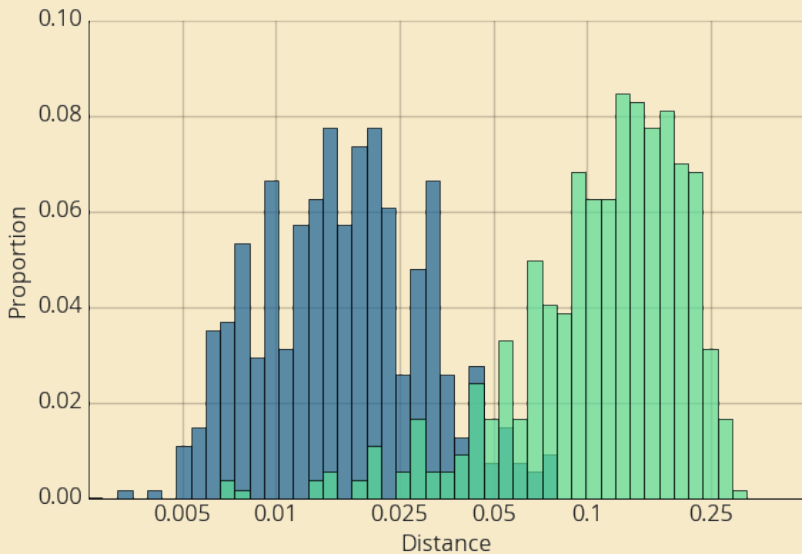
Example Embeddings



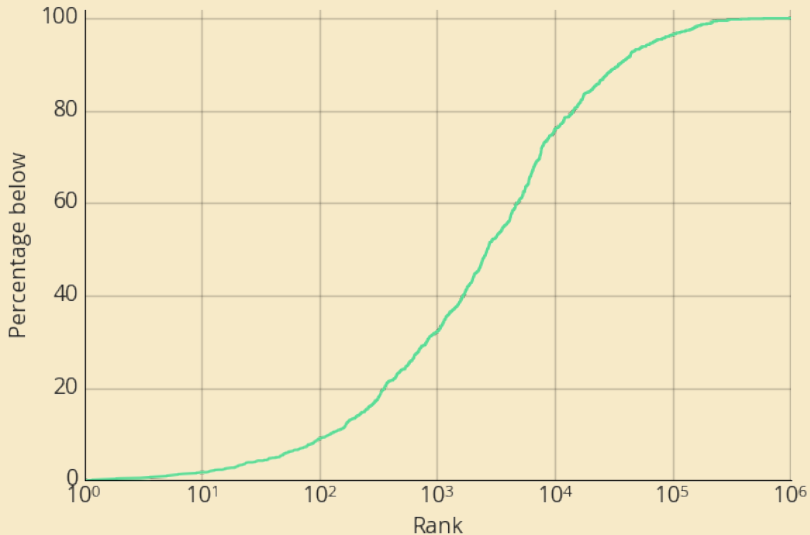
Example Attention



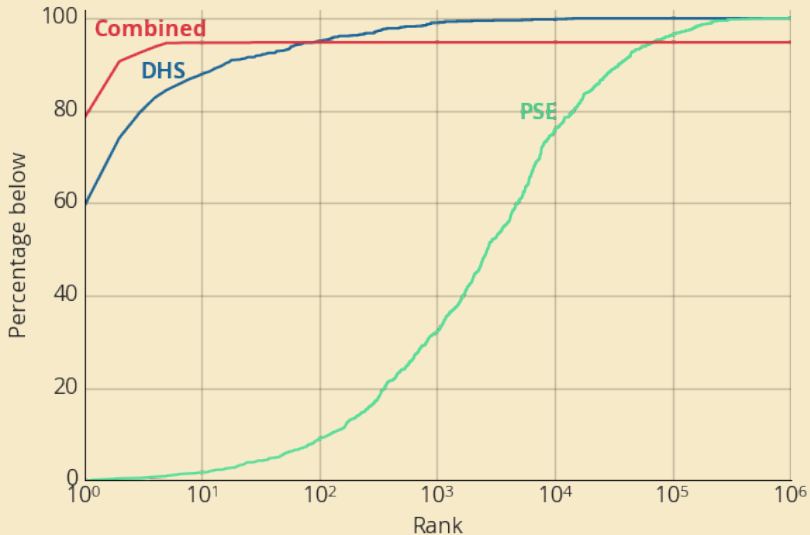
Embedding Distances



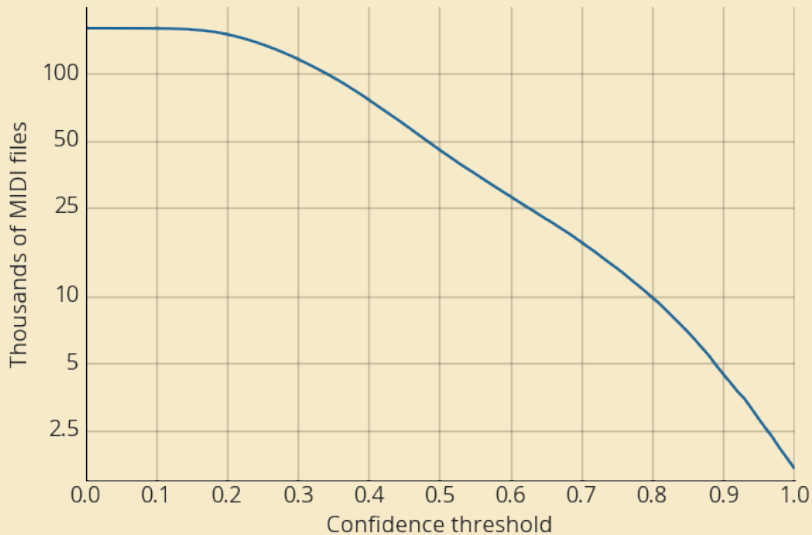
Match Ranks



Combined Match Ranks



Number of Matches



Tutorial

`http://bit.ly/lmd-tutorial`

Decoys?

Kaskade - In This Life (Mario Fabriani Remix)

Marcos Hernandez - If You Were Mine

D-Unity - Afrika

Anane - Let's Get High (Yves C Mix)

DJ Silver - Wardance

Ann Nesby - So Much Joy (Praise Party Beats)

Kim English - My Destiny (Kobbe & Austin Leeds Club Mix)

Gianluca Motta / Snap! / NG3 - Ooops Up

Dave Clarke - Protective Custody

Karizma - Ride (Original Mix)

Logic - The Warning (Claude Monnet & Torre Bros Main Mix)

Ultra Nat - Automatic (Shawn Q's Soltribe Vocal Mix)

Inaya Day - Feel It (The Down Under Mix)

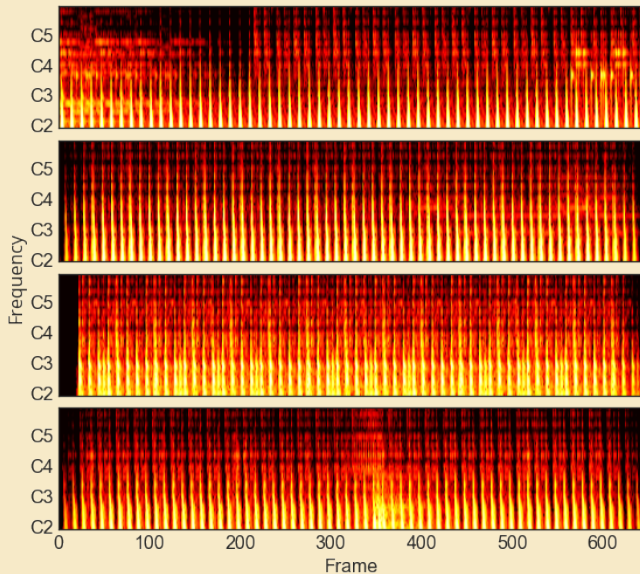
Tommyboy - Kamavtu (Cherrygirl Remix)

Michael Gray - The Weekend

Late Of The Pier - Bathroom Gurgle (Tomski*Fredboys

Tartelet Remix)

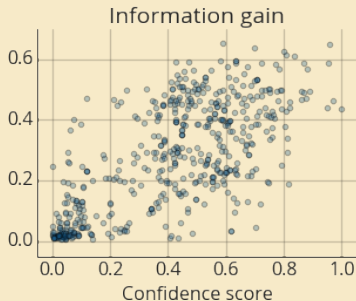
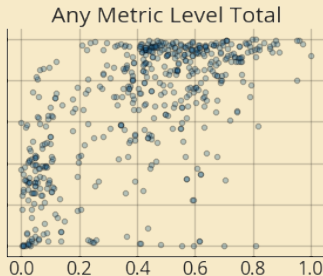
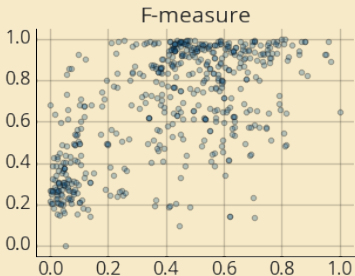
Decoys



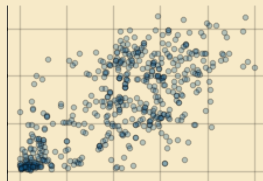
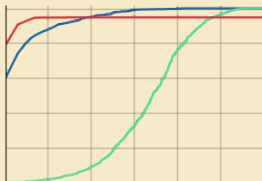
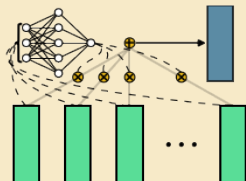
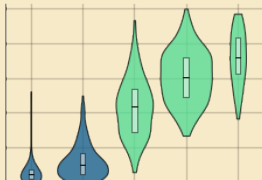
Key Reliability

Source	Score	Comparisons
MIDI, all keys	0.400	223
MIDI, C major only	0.167	146
MIDI, non-C major	0.842	77
Audio content-based	0.687	151
Human agreement	0.857	145

Beat Reliability



Summary



Questions?

