

Analisis Spektrum Akustik Tangis Bayi Penderita Penyakit Jantung Bawaan menggunakan Teknik *Dynamic Time Warping*

Hana Afifah, Iwan Yahya,* Nuryani, dan Harjana

Laboratorium Riset Akustik (iARG) Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret (UNS), Jl. Ir. Sutami 36A Kentingan, Surakarta 57126

Intisari

Paper ini berfokus pada kajian perbandingan spektral tangis bayi penderita penyakit jantung bawaan (*Congenital Heart Disease*, CHD) dan tangis bayi normal menggunakan teknik *Dynamic Time Warping* (DTW). Sebanyak sepuluh sampel tangis bayi penderita CHD dan bayi normal berusia kurang dari satu bulan direkam menggunakan Sony IC-Recorder tipe ICD-PX333. Data digital kemudian diproses menggunakan piranti lunak *Speech Filing System* (SFS) untuk mendapatkan f_0 , formants, dan nilai *Mel-Frequency Cepstral Coefficients* (MFCC) sebagai penciri spektral yang kemudian divisualkan dengan DTW. Hasil kajian menunjukkan bahwa ciri spektral tangis bayi penderita CHD mempunyai tingkat perbedaan signifikan terhadap tangis bayi normal. Di samping itu ditemukan anomali yang diduga berkaitan dengan irisan dalam bentang frekuensi ambang maupun formant tangis bayi normal dan penderita CHD.

Abstract

This paper emphasizes on acoustics spectral similarity analysis on of normal infant cry compared to infant spectrum with Congenital Heart Disease (CHD). Ten of the cries signals of infant with CHD and normal baby aged less than a month were recorded by using Sony IC-Recorder ICD-PX333. The Speech Filing System software utilized for estimating fundamental frequency, formants, and MFCCs were used as cry spectral characteristics to be visualized using Dynamic Time Warping (DTW) technique. The result shows that both group partially has significant different characteristics since anomalies occurred which is considered due to the intersection on the values of fundamental frequency and formants of the both groups.

KATA KUNCI: tangis bayi, *Congenital Heart Disease*, *Mel-Frequency Cepstral Coefficient*, *Dynamic Time Warping*

I. PENDAHULUAN

Menangis merupakan salah satu ragam vokalisasi utama pada bayi. Spektrum tangis menyajikan banyak informasi tentang kondisi bayi itu sendiri seperti lapar, sakit, mengantuk, atau bosan bahkan dapat mengindikasikan suatu penyakit [1, 2]. Para peneliti terdahulu telah melaporkan bahwa tangis bayi dapat digunakan sebagai penanda awal penyakit bawaan pada bayi bersangkutan. Salah satunya adalah penyakit jantung bawaan (*congenital heart disease*, CHD) yang berkaitan dengan kerusakan atau kelainan pada struktur jantung yang hadir pada saat lahir. Beberapa gejala CHD pada bayi berdampak pada sesak napas, sianosis, nyeri dada, sinkop, berkeringat, gumaman jantung, dan infeksi saluran pernapasan [3, 4].

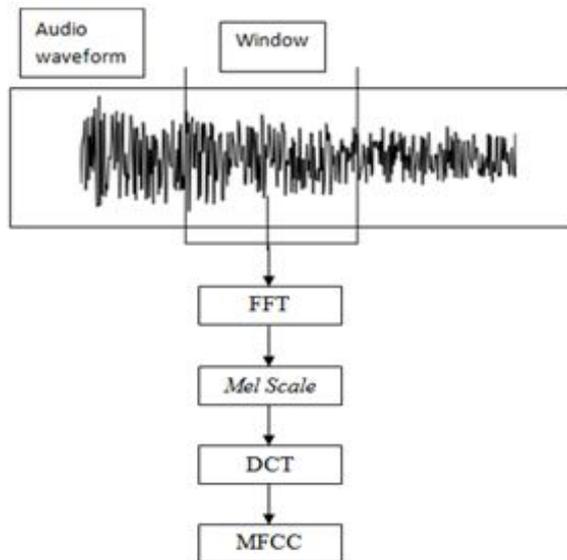
Pada umumnya para peneliti fokus pada pengukuran frekuensi ambang (f_0) pada spektrum tangis bayi sebagai parameter penanda yang berkaitan dengan gejala patologis yang dialami oleh bayi yang mereka teliti. Mereka menemukan fakta empirik yang menguatkan bahwa frekuensi ambang un-

tuk tangis bayi normal berada pada kisaran frekuensi yang berbeda dengan bayi yang menderita kelainan patologis. Fort *et al.* [5] misalnya, meneliti sinyal akustik tangis bayi dan menemukan bahwa frekuensi ambang (f_0) dari tangis bayi normal sekitar 350 Hz. Adapun Hirsberg [6] mendapatkan f_0 kisaran (400-500) Hz, sementara penderita kerusakan sistem saraf pusat mempunyai f_0 sekitar (1450-1800) Hz, bayi penderita *down syndrome* pada kisaran (60-70) Hz, penderita *micrognathia* di kisaran 4000 Hz, penderita *Paresis Nervirecurrentis* dengan f_0 sekitar 650 Hz, dan bayi penderita bronkhitis memiliki nilai f_0 sekitar 5500 Hz.

Kajian lain oleh Robb *et al.* [7] mendapatkan rerata f_0 pada bayi normal sekitar 512 Hz, sementara kajian Lind *et al.* [8] menemukan bahwa nilai f_0 bernilai 450 Hz. Bard *et al.* [9] dan Lagasse *et al.* [10] berturut-turut mendapatkan nilai f_0 bayi normal pada 515 Hz dan kisaran (250-600) Hz. Daga dan Panditrao [1] yang secara spesifik meneliti bayi normal usia kurang dari satu bulan mendapatkan f_0 sekitar 400 Hz.

Penelitian oleh Rani *et al.* [4] menemukan f_0 pada bayi normal sekitar 383 Hz, dan tiga frekuensi formant yaitu f_1 , f_2 , f_3 , secara berurutan-turut sekitar 684, 1315, 2174 Hz. Mereka menemukan pula sejumlah fakta berkaitan f_0 sebagai berikut. Bayi dengan kelainan *tetralogy of Fallot* memiliki frekuensi ambang sekitar 364-396 Hz, penderita *Ventricular Septal Defect* (VSD) pada (435-447) Hz, kemudian kisaran (435-467)

*E-MAIL: iyahya@mipa.uns.ac.id



Gambar 1: Proses akustik untuk memperoleh MFCC.

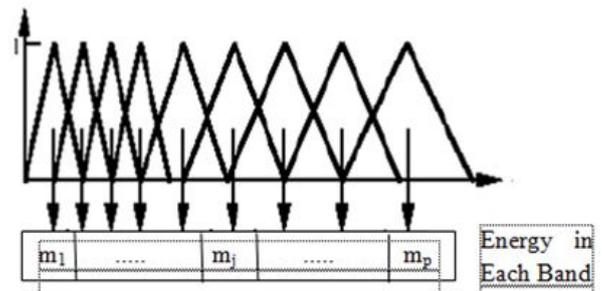
Hz untuk bayi dengan *Atrial Septal Defect* (ASD); dan (55-584) Hz untuk bayi dengan kelainan *Patent Ductus Arteriosus* (PDA). Penelitian lain oleh Michelsson *et al.* [11] diikuti Michelsson and Sirvio [12] berturut-turut untuk kasus bayi penderita *cleft palate* dan *hipotroidism* menemukan bahwa f_0 berada pada rentang (360-710) Hz dan (270-470) Hz.

Makalah ini menyajikan pendekatan berbeda dengan analisis ekstraksi ciri didasarkan atas nilai *Mel Frequency Cepstral Coefficient* (MFCC) pada sinyal tangis bayi yang kemudian dimanfaatkan untuk visualisasi lanjutan menggunakan *dynamic time warping* (DTW). Paduan antara MFCC dan DTW dalam analisis tangis bayi khususnya pada penderita gangguan jantung bawaan merupakan merupakan sumbangan pemikiran baru mengingat kajian sejenis belum banyak dipublikasikan. Kajian semisal oleh Vikas and Sharma [13] yang meneliti spektrum suara penderita penyakit parkinson maupun oleh Subali *et al.* [14] pada persoalan pengucapan 'makhraj' menunjukkan keberhasilan paduan ini.

II. METODOLOGI

Proses ekstraksi ciri

Sebanyak sepuluh sampel tangis bayi penderita CHD dan bayi normal berusia kurang dari satu bulan yang dikumpulkan dari RSUD Dr. Moewardi dan Klinik Ibu & Anak Solo Peduli. Injeksi langsung ke kulit bayi dilakukan oleh petugas medik sebagai stimulus rasa sakit. Tangis bayi kemudian direkam dengan Sony IC-Recorder tipe ICD-PX333 dan disimpan ke bentuk digital dengan ekstensi *.wav. Data digital tersebut kemudian diproses dengan aplikasi *Speech Filing System* (SFS) untuk ekstraksi f_0 , formants, dan nilai MFCC sebagai pen-ciri spektralnya. Untuk tujuan dimaksud, rekaman dari Sony



Gambar 2: Mel-frequency filterbank [18].

IC-Recorder ditapis dengan *band pass filter* pada bentang 200 Hz - 5,5 kHz dan dilakukan pencacahan ulang (re-sampling) dengan laju 11025 Hz. Proses ini dilakukan dengan perangkat lunak *Cool Edit Pro 2.1*.

Proses estimasi frekuensi ambang menggunakan analisis cepstrum, sedangkan frekuensi formants menggunakan teknik *Linier Predictive Coding* (LPC) dalam piranti lunak SFS. Kedua metode tersebut diaplikasikan ke seluruh data tangis bayi normal maupun CHD. Adapun alur proses ekstraksi MFCC disajikan dalam Gambar 1 yang mencakup empat langkah meliputi *Windowing*, *Fast Fourier Transform* (FFT), *Mel scale*, dan *Discrete Cosine Transform* (DCT).

Hamming window digunakan untuk reduksi derau di awal dan ujung frame sesuai persamaan berikut,

$$Wn(m) = 0,54 - 0,46\cos\frac{2\pi m}{Nm - 1}, 0 \leq m \leq Nm - 1 \quad (1)$$

Cuplikan hasil windowing kemudian diproses dengan algoritme FFT sehingga diperoleh data di domain frekuensi. Data dimaksud kemudian ditapis menggunakan *Mel-frequency filterbank* yang merupakan untai *triangular band pass filter* sebagaimana disajikan dalam Gambar 2. Pada tahap akhir diterapkan algoritma *Discrete Cosine Transform* (DCT) untuk menginversi data kembali ke domain waktu. Hasilnya merupakan MFCC yang dapat dihitung dengan persamaan:

$$C_j = \sqrt{\frac{2}{N}} \sum_{j=1}^p m_j \cos\left(\frac{\pi j}{N}(j - 0,5)\right) \quad (2)$$

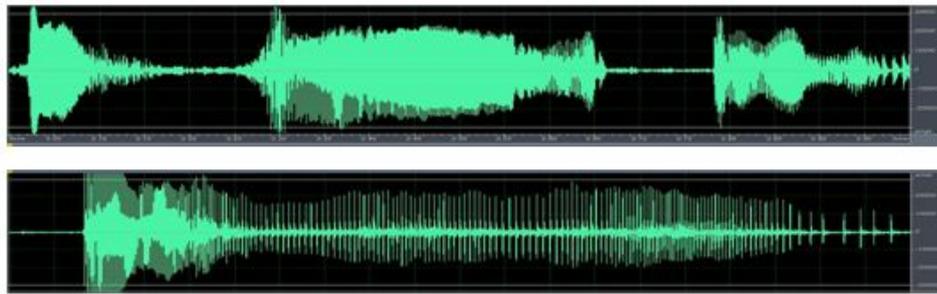
Hasil perhitungan MFCC selanjutnya digunakan sebagai masukan pada metode analisis kemiripan pola dengan DTW [15-17].

Dynamic Time warping (DTW)

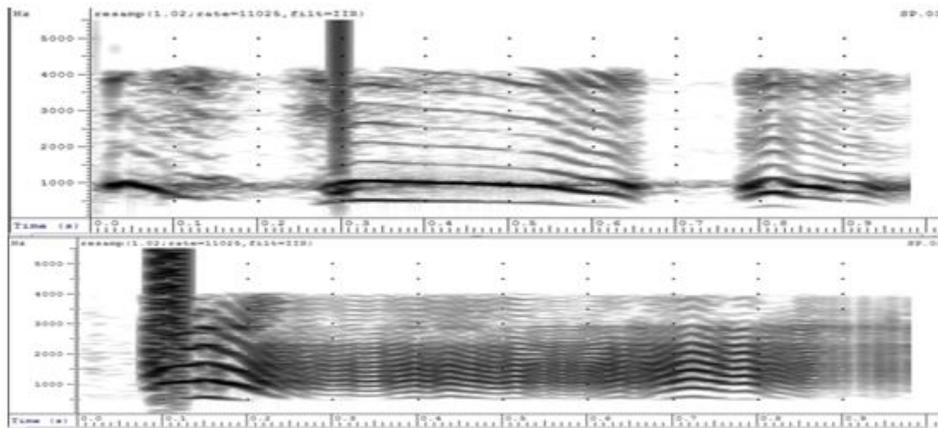
Untuk kepentingan visualisasi kemiripan ciri akustik tangis bayi, digunakan algoritma DTW yang berawal dari pembentukan jarak matriks $C \in R^{N \times M}$ mewakili jarak matriks antar pasangan X dan Y menurut hubungan berikut,

$$C_{\ell} \in R^{N \times M} : c_{i,j} = \|x_i - y_j\|, i \in [1 : N], j \in [1 : M] \quad (3)$$

Dalam hal ini sinyal tangis bayi yang telah dicuplik sepanjang 1 sekon diekstrak cirinya menggunakan perangkat lunak SFS untuk mendapatkan 16 koefisien MFCC dengan 5



Gambar 3: Sinyal tangis bayi normal (atas) dan penderita CHD (bawah) dipotong sepanjang 1 detik.



Gambar 4: Spektrogram dari cuplikan tangis bayi pada Gambar 3.

TABEL I: Perbandingan nilai frekuensi formant pada subyek yang diteliti.

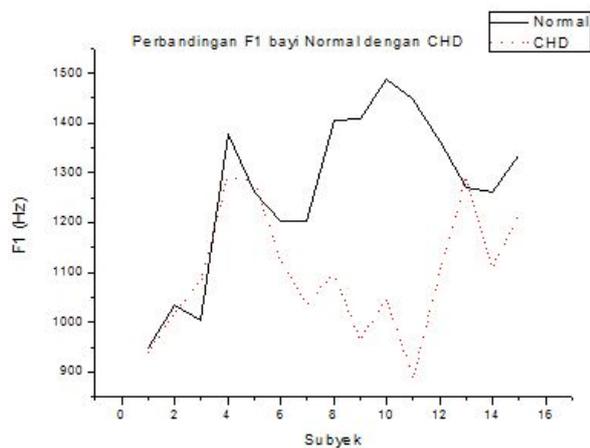
No	f_0 (Hz)		f_1 (Hz)		f_2 (Hz)		f_3 (Hz)	
	Normal	CHD	Normal	CHD	Normal	CHD	Normal	CHD
1	501	501	947	939	1706	2092	2891	2747
2	551	459	1033	1019	2793	1832	3917	2695
3	525	501	1004	1084	1670	1760	3246	2689
4	501	479	1377	1290	1917	2563	2989	4106
5	315	580	1262	1284	1921	2259	2901	3623
6	297	550	1202	1124	1778	2135	2892	3029
7	525	479	1202	1037	1713	1450	3212	2978
8	551	424	1404	1099	1708	2153	3504	3064
9	551	459	1409	965	1691	1902	3543	2904
10	393	459	1488	1046	2172	1514	3009	3151
11	551	441	1447	888	2568	1564	3616	2978
12	408	501	1363	1099	1913	1949	3419	3049
13	612	479	1269	1291	1961	2292	3128	2904
14	501	500	1261	1111	2170	2180	3128	2933
15	580	500	1334	1215	2096	2092	3103	2850

overlap FFT masing-masing pada posisi 0, 110, 220, 330, dan 440. Selanjutnya untuk kepentingan indikator kemiripan, digunakan skala keabu-abuan dengan gradasi dari warna putih terang untuk kemiripan 0% hingga hitam untuk kemiripan 100% [19, 20].

III. HASIL DAN PEMBAHASAN

Tipikal hasil perekaman tangis bayi disajikan dalam Gambar 3. Gambar tersebut merupakan contoh cuplikan untuk tangis bayi normal, sementara gambar bawah merupakan cuplikan tangis bayi yang terindikasi CHD. Analisis temporal terhadap keseluruhan tangis bayi menunjukkan bahwa subyek penderita CHD memiliki rerata durasi inspirasi sebesar 0,15 sekon, jeda 0,15 sekon dan ekspirasi 1,11 sekon. Nilai ini lebih kecil dari rerata paramater yang sama pada bayi normal yakni 0,19 sekon untuk inspirasi, jeda 0,37 sekon dan 1,27 sekon untuk ekspirasi. Adapun untuk siklus subyek penderita CHD juga memiliki rerata waktu lebih pendek sebesar 1,4 sekon sementara rerata bayi normal sebesar 1,84 sekon.

Selanjutnya analisis spektral dengan SFS menghasilkan spektrogram sebagaimana disajikan dalam Gambar 4. Hasilnya dirangkum dalam Tabel I. Cuplikan pertama hingga lima belas merupakan tangis bayi normal sementara cuplikan ke enam belas hingga tiga puluh merupakan penderita CHD. Tampak bahwa frekuensi formant dari penderita CHD memiliki kecenderungan berada pada rentang nilai yang sama dengan yang terukur pada bayi normal. Nilai f_0 untuk bayi normal berada pada kisaran (297-612) Hz, sementara penderita CHD pada (424-580) Hz. Gejala yang serupa terjadi pula pada nilai f_2 dan f_3 . Nilai f_2 untuk bayi normal adalah dalam kisaran (1,67-2,793) kHz, yang beririsan dengan nilai f_2 penderita CHD pada bentang (1,45-2,563) kHz. Adapun untuk f_3 ,



Gambar 5: Perbandingan nilai formant f_1 antara tangis bayi normal (solid) dengan penderita CHD (dotted).

bayi normal berkisar pada frekuensi (2,891-3,917) kHz, penderita CHD pada frekuensi (2,689-4,106) kHz.

Perbedaan yang cukup signifikan ditemukan pada sebagian nilai f_1 dimana bayi normal berada pada kisaran (947 - 1488) Hz sementara penderita CHD pada (888 - 1291) Hz sebagaimana disajikan dalam Gambar 5.

Cuplikan visual kemiripan ciri menggunakan DTW antara tangis bayi normal dan penderita CHD disajikan dalam Gambar 6 dan Gambar 7. Dalam hal ini masing-masing tangis dari lima bayi penderita CHD{1,2,3,4,5} dan lima bayi normal N{1,2,3,4,5} dibandingkan dengan subyek CHD{1} digunakan sebagai ciri referensi.

Gambar 6(a) merupakan uji DTW CHD{1} terhadap dirinya sendiri. Pola simetrik dengan optimal warping path berupa garis lurus serta indikator kemiripan gelap sempurna menunjukkan bahwa algoritma yang diterapkan bekerja dan mampu menganalisis tingkat kemiripan ciri tangis bayi dengan baik.

Pola sejenis untuk tangis bayi normal N{1} hingga N{5} dengan referensi tangis bayi CHD{1} disajikan dalam Gambar 7.

Gambar 6 menunjukkan bahwa terdapat konsistensi dan kemiripan yang tegas pada sebagian subyek penderita CHD

yang diteliti. Optimal warping path berbentuk garis lurus dan indikator kemiripan yang tinggi terbentuk pada sebagian ciri dari masing-masing subyek penderita CHD{2,4,5}. Sementara itu anomali teramati pada ciri subyek CHD{3} yang memiliki indikator kemiripan terendah dibandingkan dengan subyek penderita CHD yang lain.

Analisis terhadap ciri tangis bayi normal yang dibandingkan dengan CHD{1} sebagai referensi menunjukkan bahwa empat dari lima subyek tangis bayi normal menunjukkan perbedaan ciri secara konsisten. Pola yang berbeda ditemukan pada subyek N{4} dimana sebagian cirinya memiliki skala kemiripan yang tinggi dan optimal warping path berupa garis lurus sebagian.

Munculnya anomali pada hasil analisis DTW untuk subyek CHD{3} dan N{4} di atas diduga berkaitan dengan adanya irisan dalam bentang nilai antara frekuensi formant f_0 , f_1 , f_2 , dan f_3 dari subyek tangis bayi normal dan penderita CHD yang diteliti. Hasil ini dapat pula dikaitkan dengan temuan yang disajikan oleh Rani yang melakukan pengelompokan dalam kategori klinik yang lebih spesifik. Irisan dari bentak frekuensi dari hasil cuplikan terhadap subyek bayi normal dan penderita CHD menjelaskan mengapa muncul pola linear sebagian dalam analisis DTW pada Gambar 6 sekaligus menjelaskan munculnya pola anomali pada Gambar 6(c) dan 7(d).

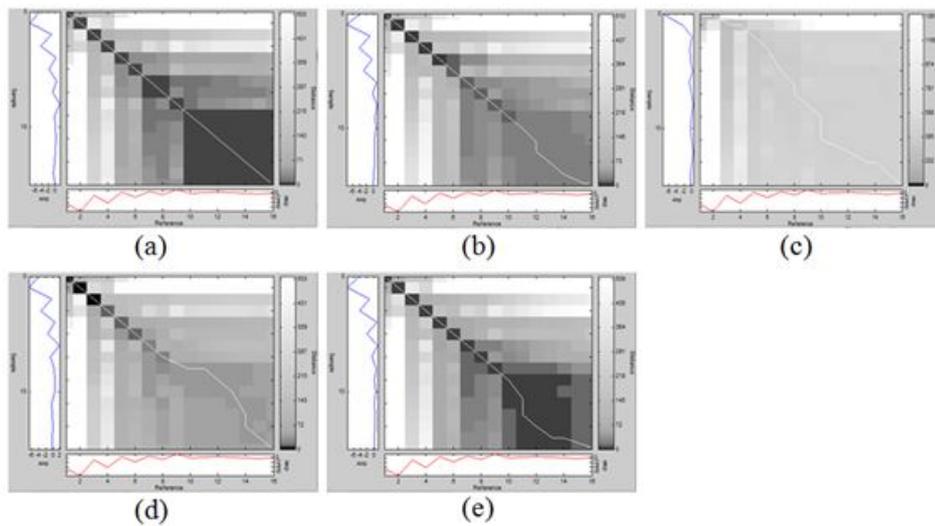
Kajian spektral menunjukkan bahwa pada kasus tertentu tangis bayi normal dan CHD mempunyai ragam spektral yang mirip satu sama lain. Namun sebuah pola menarik ditemukan pada perbandingan nilai frekuensi formant f_1 yang menunjukkan perbedaan mencolok sebagaimana disajikan pada Gambar 5. Oleh karena itu dipandang perlu kajian lanjutan dimana analisis spektral diterapkan berdasarkan sebuah prosedur klasifikasi klinik yang lebih spesifik.

IV. SIMPULAN

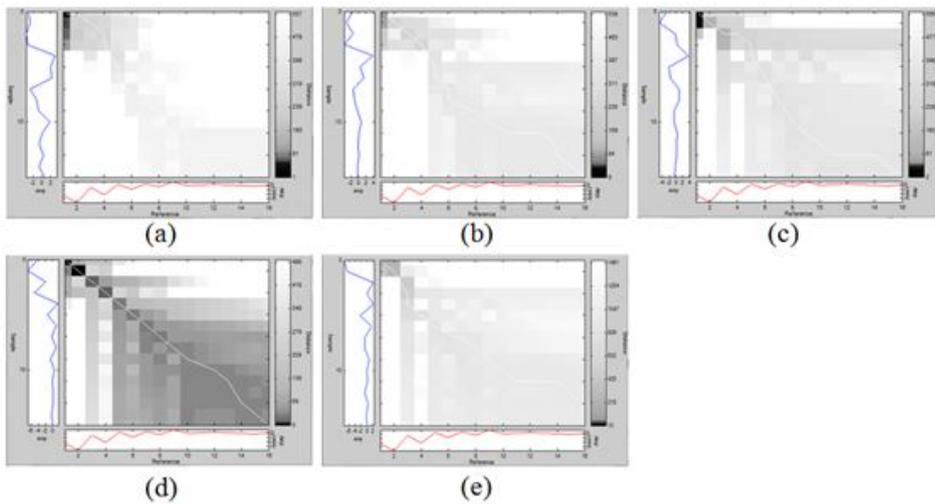
Visualisasi hasil DTW menunjukkan bahwa tangis bayi penderita CHD mempunyai tingkat perbedaan spektral yang tinggi terhadap tangis sakit bayi normal. Pada sebagian kasus ditemukan anomali yang diduga berkait dengan irisan dalam bentang frekuensi ambang maupun formant tangis bayi normal dan penderita CHD.

[1] P.R. Daga, and M.A. Panditrao, International Journal of Computer Applications (0975-8887) on Electronics, Information and Communication Engineering (ICEICE), **3**, 18-21. (2011).
 [2] B. Mijovic, et al., Schattauer, **49**, 1-5 (2010).
 [3] R. Rani, et al., International Journal of Electrical and Electronics Engineering, **1**(3), 32-37 (2012).
 [4] R. Rani, et al., Indian Journal of Computer Science and Engineering IJCSE, **3**(4), 574-582 (2012).
 [5] A. Fort, and C. Marfendi, Medical Engineering and Physics, **20**, 432-442 (1998).
 [6] J. Hirschberg, International Journal of Pediatric Otorhinolaryngology, **2**, 287-300 (1980).
 [7] M.P. Robb, and A.T. Caca, International Journal of Pediatric

Otorhinolaryngology, **32**, 57-67 (1994).
 [8] J. Lind, et al., Developmental Medicine and Child Neurology, **12**, 478-486 (1970).
 [9] D. Bard, P. Runefors, and E. Arnbjornsson, *Newborn pain cry analysis based on pitch frequency tracking*, [www.acoustic08-paris.org.] (2008).
 [10] L. Lagasse, R. Neal, B.M. Lester, Mental Retardation and Development Disabilities Research reviews, **11**(1), 83-93 (2005).
 [11] K. Michelsson, et al., Biology of the Neonate, **26**, 353-358 (1975).
 [12] K. Michelsson, and P. Sirvio, Folia Phoniatica, **26**, 40-47 (1976).
 [13] Vikas, and R.K. Sharma, Scopus, **2**, 753-759 (2015).



Gambar 6: Pola DTW tangis bayi penderita CHD berturut-turut dari (a). CHD{1} hingga (e). CHD{5} dengan referensi CHD{1}.



Gambar 7: Pola DTW tangis bayi normal berturut-turut dari (a), N{1} hingga (e) N{5} dengan referensi CHD{1}.

[14] M. Subali, M. Andriansyah, C. Sinambela, *Scopus*, **365**, 373-382 (2016).
 [15] E.A. Saputro, *Jaringan syaraf tiruan kohonen untuk analisis klinik tangis bayi*, Skripsi, Jurusan Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret, Surakarta, 2005.
 [16] M.Z.M. Ali, *et al.*, *Asphyxiated infant cry classificaton using simulink model*, Proceeding of the 8th International Colloquium on Signal Processing and its Applications, IEEE, 2012.
 [17] S. Gupta, *et al.*, *Signal & Image Processing: An International Journal (SIPIJ)*, **4**(4), 101-108 (2013).
 [18] L. Tan, and M. Karnjanadecha, *Modified Mel-Frequency Cepstrum Coefficient (ICEP)*, 2003).
 [19] P. Senin, *Dynamic Time Warping Algorithm Review (CSDL Technical Report*, 2008).
 [20] M. Muller, *Information Retrieval for Music and Motion* (Springer, Berlin Heidelberg, 2007).