*Speech enhancement* dapat didefinisikan sebagai suatu proses yang berfokus kepada pembersihan suatu sinyal suara yang telah dipengaruhi oleh sinyal *noise* tambahan. Dalam penerapannya, *speech enhancement* diperlukan untuk sinyal suara yang berasal dari tempat yang mengandung *noise* seperti tempat manufaktur. Komunikasi suara melalui telefon genggam yang dilakukan di tempat demikian akan terpengaruh oleh sinyal suara lain yang berada di tempat tersebut. Algoritma *speech enhancement* yang dapat digunakan untuk memproses suatu sinyal suara antara lain adalah *spectral subtractive algorithms*, *statistical model-based algorithms*, *subspace algorithms*, dan *binary mask algorithms* (Loizou, 2013).

Metode pertama yang dapat digunakan adalah *spectral subtractive algorithm*. *Spectral subtractive algorithm* dapat didefinisikan sebagai metode restorasi spektrum pada suatu sinyal dari suatu sinyal yang mengandung *noise*. Spektrum noise pada suatu sinyal dapat diprediksi dan diperbarui selama terdapat sinyal dan *noise* pada suatu sinyal. Dalam penerapannya, sinyal dapat ditemukan dengan cara mengurangi *noise* pada suatu sinyal yang mengandung *noise*. Namun, pada penerapannya cukup sulit untuk menghilangkan *noise* yang bersifat acak, sehingga hal yang mungkin dilakukan adalah mengurangi pengaruh dari rata-rata *noise* pada suatu spektrum sinyal. Secara umum, sinyal yang mengandung *noise* dapat dinyatakan dalam bentuk persamaan sebagai berikut:

Persamaan Persamaan umum sinyal yang mengandung noise dalam domain waktu

dengan adalah sinyal, adalah noise pada sinyal, dan adalah indeks waktu. Dalam domain frekuensi, sinyal dapat dinyatakan dalam bentuk persamaan sebagai berikut:

Persamaan Persamaan umum sinyal yang mengandung noise dalam domain frekuensi

dengan adalah sinyal dalam domain frekuensi yang telah ditransformasi menggunakan transformasi fourier, adalah *noise* yang telah ditransformasi menggunakan transformasi fourier. Dalam metode *spectral subtractive algorithm* sinyal dalam domain waktu di-*buffered* dan dibagi dalam rentang N sampel dan tiap sampel di-*window* menggunakan window Hamming atau Hanning. Setelah itu, tiap segmen yang telah di-*window* ditransformasi menggunakan transformasi Fourier disktrit (DFT). *Window* digunakan dengan tujuan untuk meringankan pengaruh diskontinuitas pada pada sampel. Persamaan windowing suatu sinyal dapat dijabarkan dalam persamaan berikut:

Persamaan Persamaan proses windowing pada sinyal dalam domain waktu

dalam domain frekuensi, persamaan windowing dapat dinyatakan sebagai:

Persamaan Persamaan windowing dalam domain frekuensi

dengan \* adalah operasi konvolusi. Secara umum, metode spectral subtraction dapat dinyatakan dalam bentuk:

Persamaan Persamaan spectral substraction

Persamaan Persamaan rata-rata spektrum noise

dengan b=1 untuk ekstraksi magnitude dan b=2 untuk ekstraksi daya. Parameter α merupakan parameter untuk menentukan jumlah sinyal noise yang akan diekstrak. Untuk pengurangan noise penuh α=1 sedangkan untuk pengurangan noise berlebih α>1 (Vaseghi, 2000).

Kekurangan dari metode *spectral subtractive* adalah metode ini menggunakan nilai parameter tetap, sehingga metode ini tidak dapat beradaptasi dengan tingkat sinyal *noise* yang ada. Optimisasi parameter pada metode ini tidak mudah jika sinyal noise yang diberikan tidak bersifat datar. Salah satu penerapan dari metode ini adalah pada ekstraksi pita ganda. Didapatkan sinyal *noise* sisa yang tidak dapat dilemahkan (Upadhyay & Karmakar, 2015).

Metode kedua yang dapat digunakan adalah metode *Statistical model-based algorithms*. Metode statistical model-based algorithms dapat didefinisikan sebagai metode untuk mengestimasi sinyal asli dengan pendekatan model statistika. Tahap pertama yang harus dilakukan pada metode ini adalah memformulasi bentuk sinyal dan *noise*. Selanjutnya, hal yang harus dilakukan adalah mengkodekan sinyal yang hendak diproses. Kemudan, sinyal yang telah dikodekan akan diklasifikasikan antara sinyal dan noise. Untuk memeriksa performa model dalam mengklasifikasikan antara sinyal dan *noise*, dapat digunakan metrics evaluasi seperti *minimum mean squared error*. Salah satu model yang dapat digunakan pada metode ini adalah menggunakan jaringan syaraf tiruan. Jaringan syaraf tiruan dapat dilatih untuk mempelajari sinyal dan *noise* lalu memberikan nilai parameter yang terbaik pada *filter* yang digunakan (Ephraim, 1992).

Metode ketiga yang dapat digunakan adalah metode *subspace algorithm*. Pada metode ini, digunakan dekomposisi *eigen value* dan transformasi Karhunen-Loeve untuk memproyeksikan suatu sinyal suara dan kemudian memisahkan antara sinyal dan noise pada suatu suara, lalu noise yang terpisahkan pada suatu sinyal suara dapat dihilangkan agar menghasilkan sinyal suara dengan kualitas yang lebih baik. Untuk meningkatkan kualitas dari sinyal yang diproses dengan menggunakan subspace algorithm, metode *perceptual feature* dapat ditambahkan untuk mengurangi distorsi pada sinyal. Berdasarkan penelitian yang telah dilakukan oleh (Surendran & Kumar, 2015) penggunaan *perceptual feature* pada metode *subspace algorithm* dapat meningkatkan nilai SNR (*Signal to Noise Ratio*) pada model.

Metode keempat yang dapat digunakan adalah metode *binary mask algorithm*. *Binary mask algorithm* dapat didefinisikan sebagai algoritma yang mengidentifikasi antara sinyal suara dan *noise* berdasarkan perbandingan antara parameter *signal to noise ratio* (SNR) pada setiap satuan TF terhadap titik ambang batas tertentu. Tahap pertama yang dilakukan pada metode ini adalah mengkonversi sinyal suara yang mengandung noise dalam bentuk vektor ke dalam bentuk matriks. Tahap ini bertujuan untuk mengubah suatu sinyal yang awal nya dalam bentuk deret menjadi bentuk matriks. Pada tahap ini, sinyal dapat dibagi menjadi matriks dengan ukuran 256 baris yang diambil setiap 32 sample dan setiap 32 sample dilakukan *windowing* dengan menggunakan window hanning. Tahap kedua yang dilakukan adalah menerapkan algoritma FFT terhadap sinyal yang telah diubah dalam bentuk matriks. Tahap ketiga yang dilakukan adalah menghitung nilai SNR pada sinyal yang telah ditransformasi menggunakan transformasi fourier. Tahap keempat adalah inisiasi batas ambang minimal dari nilai SNR untuk diteruskan menjadi nilai satu. Pada tahap ini, sinyal yang diteruskan menjadi nilai satu, dianggap menjadi sinyal suara jernih sedangkan sinyal yang diteruskan menjadi nol, dianggap menjadi sinyal *noise*. Tahap kelima adalah, menerapkan Inverse Fast Fourier Transform (IFFT), tahap ini bertujuan untuk mengubah sinyal dalam domain frekuensi ke dalam domain waktu. Tahap terakhir adalah mengubah sinyal dalam bentuk matriks menjadi vektor, tahap ini bertujuan untuk mengembalikan sinyal dalam bentuk matriks saat diproses menjadi bentuk sinyal kontinu (Patil & Mane, n.d.).

DAFTAR PUSTAKA

Ephraim, Y. (1992). *Statistical-Model-Based Speech Enhancement Systems*.

Loizou, P. C. (2013). *Speech Enhancement : Theory and Practice, Second Edition* (Second Edition). Taylor & Francis.

Patil, P. v, & Mane, V. A. (n.d.). Speech Enhancement by using Ideal Binary Mask. *International Journal of Engineering and Techniques*, *3*. http://www.ijetjournal.org

Surendran, S., & Kumar, T. K. (2015). Perceptual Subspace Speech Enhancement with Variance Normalization. *Procedia Computer Science*, *54*, 818–828. https://doi.org/10.1016/j.procs.2015.06.096

Upadhyay, N., & Karmakar, A. (2015). Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study. *Procedia Computer Science*, *54*, 574–584. https://doi.org/10.1016/j.procs.2015.06.066

Vaseghi, S. v. (2000). *Advanced digital signal processing and noise reduction*. John Wiley.