Armenian Journal of Physics, 2026, vol. 19, issue 1, pp. 86-92

Time–Frequency Feature Analysis of Speech Signals Using Wavelet Scalograms and Gabor Transform.

M.A. Ghukasyan[1], E.R. Sivolenko[1, 2]

[1]*Yerevan State University*, *1 Alex Manoukian str., 0025, Yerevan, Republic of Armenia*
[2]*Russian-Armenian University*, *123 Hovsep Emin str., 0056, Yerevan, Republic of Armenia*

Email: mariam.ghukasyan7@edu.ysu.am

(Received: January 30, 2026; Revised: February 15, 2026; Accepted: March 2, 2026)

Abstract. Speech signals are inherently non-stationary and require effective time–frequency analysis techniques to capture their dynamic characteristics. This study investigates the extraction and analysis of time–frequency features from speech signals using wavelet scalograms and the Gabor transform. The wavelet scalogram provides a multi-resolution representation of the signal, enabling detailed analysis of both high- and low-frequency components over time. In contrast, the Gabor transform offers a localized time–frequency representation using Gaussian-windowed sinusoids, which is well suited for analyzing spectral structures in speech. In this work, speech signals are processed using both approaches to obtain their respective time–frequency coefficients. The resulting representations are then analyzed to examine differences in resolution, feature localization, and their effectiveness in capturing speech characteristics. Experimental results demonstrate that the wavelet scalogram provides better adaptability to transient features, while the Gabor transform offers consistent resolution across the time–frequency plane. The comparative analysis highlights the strengths of each method and their potential applications in speech analysis, feature extraction, and signal characterization.

Keywords: Speech Signal Analysis; Time–Frequency Analysis; Wavelet Scalogram; Gabor Transform; Feature Extraction; Non-Stationary Signals; Time–Frequency Representation; Speech Processing.

DOI: 10.54503/18291171-2026.19.1-86

1. Introduction

The analysis of signals whose spectral content varies over time is a central problem in modern signal processing, harmonic analysis, and applied mathematics. Classical Fourier analysis provides information about the global frequency content of a signal but does not reveal how these frequencies evolve in time. To overcome this limitation, a variety of time–frequency and time–scale representations have been developed that allow the simultaneous localization of signal structures in both domains.

Among the most important approaches are methods based on windowed Fourier analysis and wavelet analysis. Two representative tools within these frameworks are the Gabor transform, which produces Gabor coefficients, and the continuous wavelet transform, whose energy representation is known as the scalogram. Both techniques provide localized representations of signals and are widely used in fields such as speech processing, biomedical signal analysis, radar, seismic analysis, and image processing.

Although scalograms and Gabor coefficients serve similar purposes in visualizing and analyzing nonstationary signals, they arise from different mathematical constructions. The Gabor transform relies on translations and modulations of a fixed Gaussian window, producing a uniform tiling of the time–frequency plane. In contrast, the wavelet transform employs dilations and translations of a mother wavelet, resulting in a multiresolution representation that adapts its temporal and spectral resolution across scales.

Understanding the conceptual and mathematical differences between these two representations is important for selecting appropriate analysis methods in practical applications. This article introduces the theoretical background of scalograms and Gabor coefficients and provides a technical comparison of their underlying transforms, resolution properties, and typical areas of application.

2. Scalogram and Gabor Coefficients: Technical Comparison

A scalogram is the energy representation of the continuous wavelet transform of a signal. The term scalogram refers to the squared magnitude of the wavelet coefficients and provides a time–scale representation of the signal. In signal processing literature it is often called the wavelet scalogram or CWT energy density representation.

Mathematically, the continuous wavelet transform of a signal $x(t)$with respect to a mother wavelet $\psi(t)$ is

$$W_x(a,b) = \int_{-\infty}^{\infty} x(t)\, \psi^* \left(\frac{t-b}{a}\right) dt \tag{1}$$

where a denotes the scale parameter and b denotes the time shift. The scalogram is defined as the squared magnitude of these coefficients

$$S(a,b) = \mid W_x(a,b) \mid^2 \tag{2}$$

The scalogram, therefore, describes how the energy of the signal is distributed over time and scale. Because scale is inversely related to frequency, the scalogram provides a non-linear time–frequency representation.

A Gabor coefficient is a coefficient obtained from the Gabor transform, which can be interpreted as a short-time Fourier transform using a Gaussian window. In harmonic analysis, these coefficients are also called Gabor frame coefficients.

The Gabor transform of a signal $x(t)$ with a window function $g(t)$ is

$$G_x(t,\omega) = \int_{-\infty}^{\infty} x(\tau) g(\tau - t)\, e^{-i\omega\tau} d\tau \tag{3}$$

where t represents time and ω represents angular frequency. The resulting complex values $G_x(t,\omega)$ are the Gabor coefficients and describe the local frequency content of the signal around time t.

The scalogram represents signal energy as a function of time and scale, while the Gabor representation describes signal content as a function of time and frequency. Although both methods provide time–frequency information they rely on fundamentally different analysis functions. Wavelet analysis employs scaled and translated wavelets, which change their temporal width depending on the scale parameter. This produces a multi-resolution representation. High frequencies are analyzed with short windows that provide good time localization, whereas low frequencies are analyzed with longer windows that provide improved frequency resolution.

The Gabor transform uses a Gaussian window of fixed width. Consequently, the time–frequency resolution remains constant across the entire representation. This property produces a uniform lattice in the time–frequency plane.

Scalograms originate from the theory of wavelets and multiresolution analysis, whereas Gabor coefficients originate from time–frequency analysis and frame theory. Wavelet transforms employ logarithmic frequency sampling due to the scaling operation. This leads to denser sampling at low frequencies and coarser sampling at high frequencies. The Gabor transform uses linear frequency sampling, which produces equally spaced frequency bins similar to those of the Fourier transform. Because of this difference the wavelet scalogram adapts to signals containing short transient structures, while the Gabor representation is often preferred when signals exhibit relatively stationary oscillatory components.

3. LabVIEW implementation of the Continuous Wavelet Transform (CWT) using an Analytic Wavelet Transform (AWT)

The program implements the Analytic Wavelet Transform (AWT) of a signal. The goal of the algorithm is to convert a one-dimensional time signal into a time scale representation that shows how the frequency content of the signal changes with time. The result of the computation is a matrix of complex wavelet coefficients that describe the correlation between the signal and scaled versions of a mother wavelet.

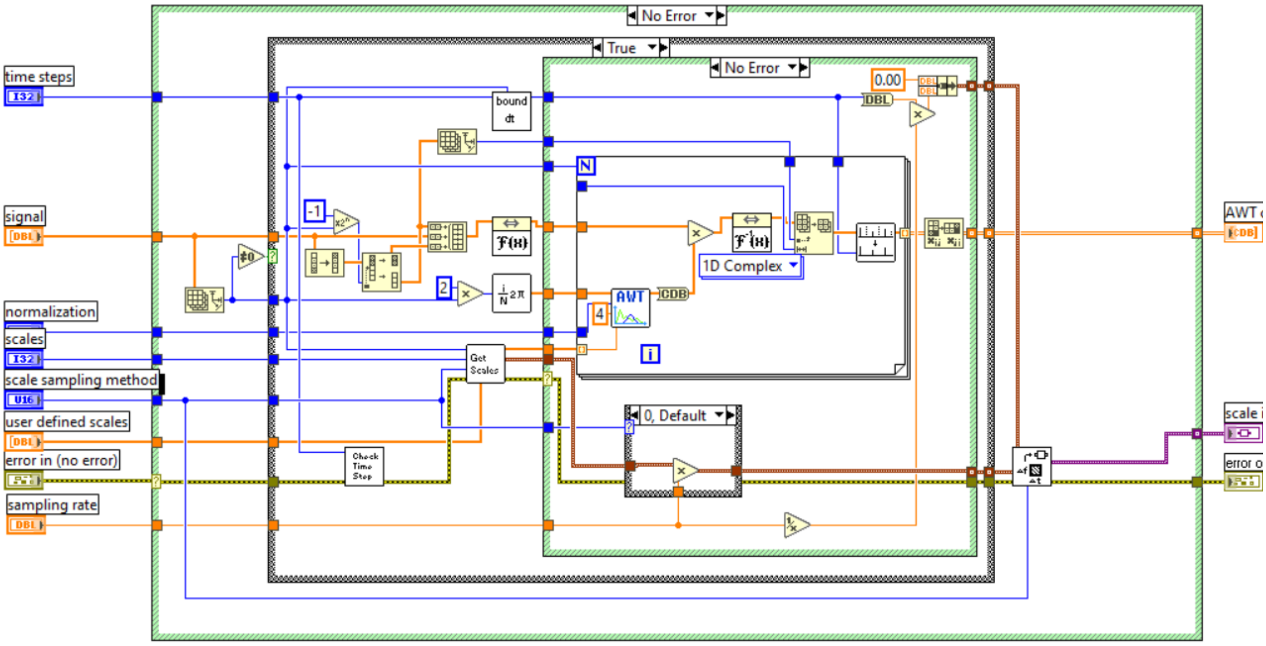

Fig. 1. LabVIEW implementation of the wavelet signal processing.

The inputs of the program are the signal, the number of time samples, the sampling rate, the number of scales, the normalization option, the scale sampling method, optional user defined scales, and the error cluster used by LabVIEW for error propagation. The outputs are the complex wavelet coefficients, the scale index and the error output.

The first part of the program performs preprocessing of the signal parameters. The number of samples is denoted by N. If the sampling rate is f_s, then the sampling interval is

$$\Delta t = \frac{1}{f_s} \qquad (4)$$

A frequency grid is constructed for the Fourier transform. The discrete angular frequencies used by the FFT are

$$\omega_k = \frac{2\pi k}{N\Delta t} \qquad (5)$$

for $k = 0,1,2, \ldots, N-1$. These values define the spectral locations where the signal spectrum will be evaluated.

The next stage determines the wavelet scales. The program contains a block that generates the scale vector according to the selected method. If the user provides the scales directly, they are used without modification. If automatic generation is selected, scales may be created using linear spacing

$$a_i = a_{\min} + i\Delta a \quad (6)$$

or logarithmic spacing, which is common in wavelet analysis

$$a_i = a_0 2^{i/s} \quad (7)$$

where s is the number of voices per octave. Each scale corresponds to a specific analysis frequency. The approximate relationship between scale and frequency is

$$f = \frac{f_c}{a} \quad (8)$$

where f_c is the center frequency of the mother wavelet. The main computation occurs inside a loop that iterates over all scales. Before entering the loop, the program computes the Fourier transform of the signal

$$X(\omega) = \mathcal{F}\{x(t)\} \quad (9)$$

This converts the time signal into its spectral representation. The wavelet transform can be efficiently computed in the frequency domain using the convolution theorem.
For every scale $a$the algorithm generates a scaled version of the mother wavelet. If $\psi(t)$is the mother wavelet then the scaled wavelet is

$$\psi_a(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t}{a}\right) \quad (10)$$

The factor $1/\sqrt{a}$ ensures energy normalization across scales. The Fourier transform of the scaled wavelet is

$$\Psi_a(\omega) = \sqrt{a}\,\Psi(a\omega) \quad (11)$$

The convolution between the signal and the scaled wavelet in the time domain is equivalent to multiplication in the frequency domain. The program multiplies the signal spectrum by the complex conjugate of the scaled wavelet spectrum

$$X(\omega)\Psi_a^*(\omega) \quad (12)$$

This operation computes the spectral product corresponding to the wavelet correlation. The inverse Fourier transform is then applied to obtain the wavelet coefficients in the time domain

$$W_x(a,b) = \mathcal{F}^{-1}[X(\omega)\Psi_a^*(\omega)] \quad (13)$$

These coefficients are complex numbers. The real and imaginary parts represent the phase-sensitive correlation between the signal and the wavelet. The magnitude of the coefficient is

$$|\ W_x(a,b)\ | \quad (14)$$

and the energy representation is

$$|\ W_x(a,b)\ |^2 \quad (15)$$

For each scale the resulting vector of coefficients corresponds to all time positions of the signal. These vectors are assembled into a two-dimensional matrix where rows represent scales and columns represent time samples.

The mathematical definition of the continuous wavelet transform is

$$W_x(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(t)\ \psi^*\left(\frac{t-b}{a}\right)dt \quad (16)$$

In this equation $x(t)$ is the signal, $\psi(t)$ is the mother wavelet, a is the scale parameter and b is the translation parameter that represents time shift. The complex conjugate of the wavelet ensures correct correlation with the signal.

The result of the algorithm is the analytic wavelet transform coefficient matrix. Each element describes how strongly the signal resembles the wavelet at a specific time and scale. Large coefficients indicate the presence of signal components whose frequency corresponds to that scale.
The wavelet transform is particularly useful for nonstationary signals because it provides adaptive time frequency resolution. High frequencies correspond to small scales and therefore provide good time localization. Low frequencies correspond to large scales and therefore provide better frequency resolution.

The output matrix can be visualized as a scalogram where the horizontal axis represents time, the vertical axis represents scale or frequency, and the color intensity represents the magnitude of the wavelet coefficients. This representation allows detailed analysis of transient features, frequency modulation and localized signal events.

4. Key results

The figure presents two different approaches for time–frequency analysis of the same signal. The left side illustrates the processing based on the Gabor transform, while the right side shows the wavelet transform representation through a scalogram. Although both methods analyze how signal frequencies evolve over time, they differ in mathematical formulation, resolution properties, and the structure of the resulting representations.

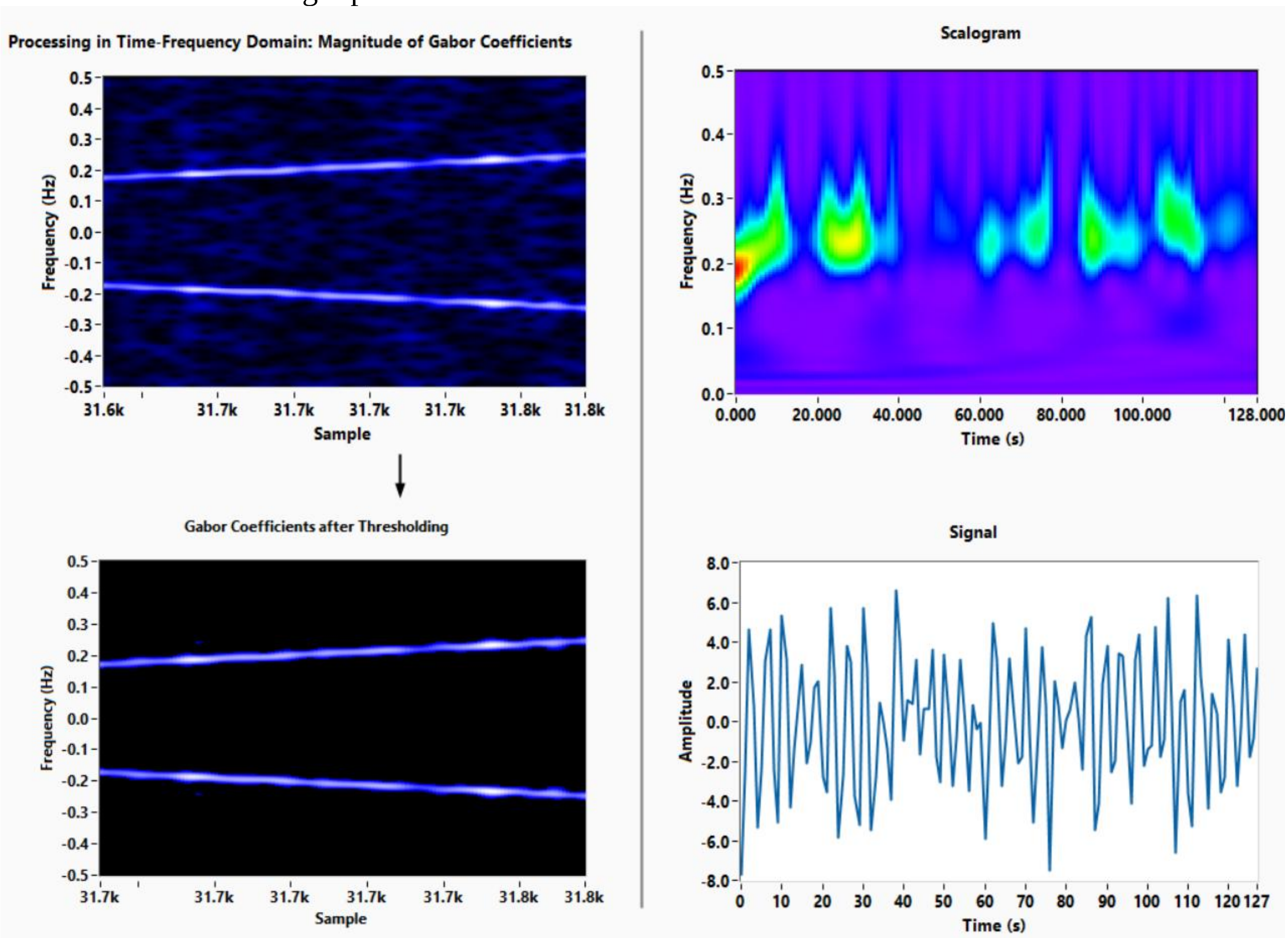

Fig. 2. Gabor transformation outcome from left and wavelet transform from right.

The left part of the figure 2 shows the magnitude of the Gabor coefficients and the result after applying thresholding.

The top-left panel represents the magnitude of the coefficients obtained from the Gabor transform, which is a form of the Short-Time Fourier Transform (STFT) using a Gaussian window. In the plot, two clear frequency ridges appear around approximately 0.2Hz and −0.2Hz. These ridges correspond to dominant oscillatory components in the signal. Because the Gabor transform uses a fixed window size, the time and frequency resolution remain constant across the entire representation.

Bottom-left panel a threshold has been applied to the coefficient magnitude. Small coefficients associated with noise or weak components are removed, leaving only the dominant spectral ridges. The result highlights the principal frequency trajectories of the signal and suppresses background noise.

The right side of the figure 2 shows the scalogram and the corresponding time-domain signal. Top-right, the scalogram, represents the energy distribution of the signal in time and scale, obtained from the Continuous Wavelet Transform (CWT). The horizontal axis represents time, while the vertical axis represents frequency (or scale). Color intensity corresponds to signal energy at each time–frequency location.

Unlike the Gabor representation, the wavelet transform provides adaptive resolution. High-frequency components are represented with better time localization, while low-frequency components have better frequency resolution. The scalogram therefore, shows broader energy regions rather than thin ridges, reflecting the variable resolution of the wavelet analysis.

Bottom-right panel: shows the original signal amplitude as a function of time. The oscillatory behavior observed here corresponds to the dominant frequency components detected in both time–frequency representations.

Key Differences Between the Left and Right Parts

Aspect	Left Side (Gabor Transform)	Right Side (Wavelet Transform)
Method	Short-Time Fourier Transform with Gaussian window	Continuous Wavelet Transform
Output representation	Magnitude of Gabor coefficients	Scalogram
Resolution	Constant time–frequency resolution	Adaptive resolution
Frequency representation	Direct frequency axis	Scale or pseudo-frequency
Feature appearance	Thin frequency ridges	Energy blobs or patches
Noise suppression	Thresholding applied	Energy naturally localized
Best suited for	Stationary or slowly varying signals	Nonstationary or transient signals

Conclusion

The left side clearly identifies precise frequency trajectories using the Gabor transform. The constant resolution allows narrow frequency lines to be tracked effectively, especially after thresholding removes noise.

The right side reveals how the signal energy evolves over time using the wavelet transform. The scalogram emphasizes transient behavior and localized bursts of energy, which are less sharply defined in the Gabor representation.

Together, these representations demonstrate two complementary approaches to time–frequency signal analysis: the Gabor transform offers precise frequency localization, while the wavelet transform provides adaptive time–frequency resolution that is better suited for analyzing nonstationary signals.

References

[1] D. Gabor, “Theory of Communication,” *Journal of the Institution of Electrical Engineers*, vol. 93, no. 26, pp. 429–457, 1946.

[2] J. B. Allen and L. R. Rabiner, “A Unified Approach to Short-Time Fourier Analysis and Synthesis,” *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1558–1564, 1977.

[3] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA, USA: SIAM, 1992.

[4] L. Cohen, *Time–Frequency Analysis*, Upper Saddle River, NJ, USA: Prentice Hall, 1995.

[5] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed., San Diego, CA, USA: Academic Press, 1999.

[6] P. Flandrin, *Time–Frequency/Time–Scale Analysis*, San Diego, CA, USA: Academic Press, 1999.

[7] C. Torrence and G. P. Compo, “A Practical Guide to Wavelet Analysis,” *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61–78, 1998.

[8] B. Boashash, *Time Frequency Signal Analysis and Processing*, Oxford, UK: Elsevier, 2003.