

**Talat Zahra 52734**

Digital signal processing project

AI +DSP RELATED PROJECT

# Smart Voice Activity Detection (VAD) Demo Synthetic Audio

**clc; clear; close all;**

# Step 1: Generate synthetic audio (speech-like signal)

**fs = 8000; % Sampling frequency (8 kHz) t = 0:1/fs:2; % 2 seconds duration**

**audio = 0.6\*sin(2\*pi\*300\*t) + 0.4\*sin(2\*pi\*600\*t); % sum of two sinusoids**

**(simulated speech)**

**audio = audio(:); % column vector audio = audio / max(abs(audio)); % normalize amplitude**

# Step 2: Frame the signal

**frame\_duration = 0.02; % 20 ms frames**

**frame\_len = round(fs \* frame\_duration); frame\_shift = round(frame\_len / 2); % 50% overlap**

**num\_frames = floor((length(audio) - frame\_len) / frame\_shift) + 1;**

**frames = zeros(frame\_len, num\_frames);**

**for i = 1:num\_frames**

**idx = (i-1)\*frame\_shift + 1 : (i-1)\*frame\_shift + frame\_len;**

**frames(:,i) = audio(idx); end**

# Step 3: Compute short-time energy and zerocrossing rate

**energy = sum(frames.^2); % Energy per frame zcr = sum(abs(diff(sign(frames))) > 0); % Zero-crossing rate per frame**

# Step 4: Voice Activity Detection (VAD)

**energy\_thresh = 0.01; % Energy threshold zcr\_thresh = 20; % Zero-crossing threshold vad = (energy > energy\_thresh) & (zcr < zcr\_thresh);**

# Step 5: Plot results

**time\_audio = (0:length(audio)-1) / fs; figure('Name','Voice Activity Detection Demo','NumberTitle','off');**

**subplot(3,1,1); plot(time\_audio, audio); title('Synthetic Speech Signal'); xlabel('Time (seconds)'); ylabel('Amplitude'); grid on;**

**subplot(3,1,2); stem(vad, 'filled');**

**title('VAD Output (1 = Speech, 0 = Silence)');**

**xlabel('Frame Number'); ylabel('VAD Decision'); ylim([-0.1 1.1]); grid on;**

**subplot(3,1,3);**

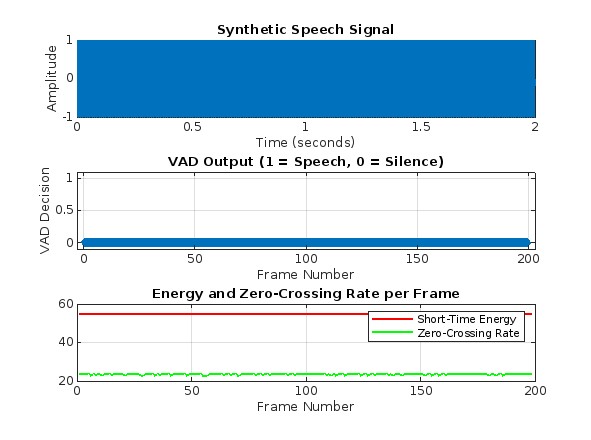
**plot(energy, 'r', 'LineWidth', 1.5);**

**hold on;**

**plot(zcr, 'g', 'LineWidth', 1.5); legend('Short-Time Energy','Zero-Crossing Rate'); title('Energy and Zero-Crossing Rate per Frame');**

**xlabel('Frame Number');**

**grid on;**



CODE SIMULATION MATLAB

|  |
| --- |
| %% Smart Voice Activity Detection (VAD) Demo - Synthetic Audio  clc; clear; close all;  %% Step 1: Generate synthetic audio (speech-like signal)  fs = 8000; % Sampling frequency (8 kHz)  t = 0:1/fs:2; % 2 seconds duration  audio = 0.6\*sin(2\*pi\*300\*t) + 0.4\*sin(2\*pi\*600\*t); % sum of two sinusoids (simulated speech)  audio = audio(:); % column vector  audio = audio / max(abs(audio)); % normalize amplitude  %% Step 2: Frame the signal  frame\_duration = 0.02; % 20 ms frames  frame\_len = round(fs \* frame\_duration);  frame\_shift = round(frame\_len / 2); % 50% overlap  num\_frames = floor((length(audio) - frame\_len) / frame\_shift) + 1;  frames = zeros(frame\_len, num\_frames);  for i = 1:num\_frames  idx = (i-1)\*frame\_shift + 1 : (i-1)\*frame\_shift + frame\_len;  frames(:,i) = audio(idx);  end  %% Step 3: Compute short-time energy and zero-crossing rate  energy = sum(frames.^2); % Energy per frame  zcr = sum(abs(diff(sign(frames))) > 0); % Zero-crossing rate per frame  %% Step 4: Voice Activity Detection (VAD)  energy\_thresh = 0.01; % Energy threshold  zcr\_thresh = 20; % Zero-crossing threshold  vad = (energy > energy\_thresh) & (zcr < zcr\_thresh);  %% Step 5: Plot results  time\_audio = (0:length(audio)-1) / fs;  figure('Name','Voice Activity Detection Demo','NumberTitle','off');  subplot(3,1,1);  plot(time\_audio, audio);  title('Synthetic Speech Signal');  xlabel('Time (seconds)');  ylabel('Amplitude');  grid on;  subplot(3,1,2);  stem(vad, 'filled');  title('VAD Output (1 = Speech, 0 = Silence)');  xlabel('Frame Number');  ylabel('VAD Decision');  ylim([-0.1 1.1]);  grid on;  subplot(3,1,3);  plot(energy, 'r', 'LineWidth', 1.5);  hold on;  plot(zcr, 'g', 'LineWidth', 1.5);  legend('Short-Time Energy','Zero-Crossing Rate');  title('Energy and Zero-Crossing Rate per Frame');  xlabel('Frame Number');  grid on; |

OUTPUT

