IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 14-16

Enhanced Features for Stuttering Speech Signal

Ms. S.Poornima
Assistant Professor Department of computer science
VLB Janakiammal College of Arts and Science
Coimbatore – 641042

Abstract: *Speech processing is an interesting area for understanding different way of communication in signal processing. Even though various features have reported in the literature of stuttering speech signal analysis, Histogram Oriented Gradient (HOG) and Spectral Histogram Oriented Gradient (SHOG) methods plays vital role in detection of stuttered word. These features are fed to classifier to segregate normal words from abnormal word. This research paper explain the process of these two enhanced features briefly.*
Key words: *HOG, SHOG*

I. INTRODUCTION

This research work consists of three folds to detect stuttering speech segments. In the first phase, basic features such as, zero crossing, short time energy, auto correlation and average magnitude have been calculated. A threshold value based on these features has been defined to segregate words and number of gaps with in a word. Stuttered words have been identified based on the number of gaps and other metrics. In second phase, simple features such as number of frames, average and total energy, auto correlation and number of zero crossing were calculated for both normal and stuttered words. These metrics helps to differentiate stuttered words from normal words. New features such as Histogram Oriented Gradient (HOG) and Spectral based Histogram Oriented Gradient (SHOG) were calculated for both normal and disordered signal.

Histogram of oriented gradient (HOG) feature

Histogram of oriented gradient (HOG) is mainly used for detect the object in image processing and computer vision and it is a gradient based feature. In June 2005 HOG is described by Navneet Dalal and Bill Triggs [1]. This feature counts the gradient orientation occurrences in local part of an image and it is behaved like edge detection. In image processing Histogram of oriented gradient is giving good performance in face reorganization.

Algorithm of Histogram of Oriented gradients (HOGs)

HOG algorithm is implemented by few steps such as gradient computation, orientation binning, descriptors block and block normalization. Figure 1 shows steps of HOG algorithm

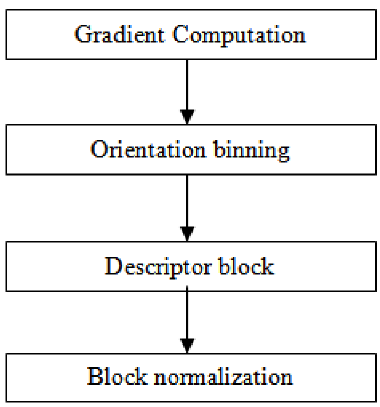

Figure 1. Steps of HOG algorithm

Theoretically first the input image is divided into connected or spatial region switch is called as Cells and calculated histogram of edge orientation for each cell. Depending on gradient 'unsigned' or 'signed' option the computing histograms are spread over like 0-180º or 0-360º and it is counted and normalized for better invariance to illumination. This process is done by accumulating the measure of a local histogram 'energy' over the larger connected or spatial region and using the results to normalize all cells in the block. Finally these normalized descriptor block is referred as "Histogram oriented gradient (HOG)descriptors" [2].

Step 1: In HOG algorithm first step is calculated gradient values from input which is used some basic method to apply the 1D centre point discrete derivative mask in both horizontal and vertical direction. The following kernel is used to filter the gray scale input

$$D_x = [-1\ 0\ 1] \text{ and } D_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad (1)$$

Where x and y is a derivatives which can be find out by using convolution operations are $I_x$andI_y.

$$I_X = I * D_X \quad (2)$$

$$I_Y = I * D_Y \quad (3)$$

A magnitude of the gradient is calculated by,

$$|G| = \sqrt{I_X^2 + I_Y^2} \quad (4)$$

The orientation of gradient is given by,

$$\theta = \arctan \frac{I_Y}{I_X} \quad (5)$$

Step 2: In the second step the cell histogram is created. Based on gradient "unsigned" or "signed" signal a histogram channels are extend 0 to 180 or 0 to 360 degrees.

Step 3: Each created cells are locally normalized and grouping together and making larger which is spatially connected with blocks. After the component of normalized cell histogram is described by HOG descriptor form each block**.** These blocks are overlap and each cell contributes with final descriptors.

Step 4: Finally a block normalization happen by using different methods. Some normalization factors are

$$L2-norm: f = \frac{V}{\sqrt{\|V\|_2^2 + e^2}}$$

$$L1-norm: f = \frac{V}{\|V\|_1 + e}$$

$$L1-sqrt: f = \sqrt{\frac{V}{\|V\|_1 + e}}$$

Where V is a non-normalized vector which include all histogram of block $\|v\|_k$, k=1,2 and e is a some small constant .

Spectral Histogram of Oriented Gradient (SHOG) Feature Extraction

Spectral Histogram of oriented gradient (SHOG) is computed by Histogram of oriented gradient (HOG) feature. A. Muthamizh selvan and R. Rajesh computed Spectral Histogram of Oriented Gradient (SHOG) using MFCC. Mel-Frequency Cepstal coeffient is computed by using discrete cosine transform (DCT) and it is used to getting spectral temporal dynamic feature from characteristics of short time spectrum. Spatial feature is offered solid overlapping description of spectral coefficient. Information of spectral local portion is called spatial feature or region. The spatial region is obtained accurately using Histogram of oriented gradient feature and the region is encoded by spectral histogram of oriented gradient method. Few steps are used to extract the SHOG from speech signal. Figure 2 shows the Block diagram of computation of SHOGs,

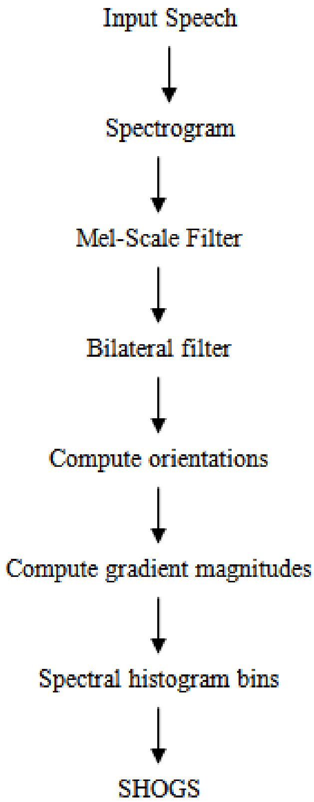

Figure 2. Block diagram of SHOGs

Step 1: From the input speech signal short term frames of time localized frequencies (t,f) and its power spectral densities p(t,f) are extracted by using spectrogram.
Step 2: The bilateral filters applied for develop and remove the noisy components from output mel scale filtered spectrum speech signal. From the signal domain and range the bilateral filter is get the distance clear notation which concerned for any signal function.
Step 3: From the mel scale filtered spectrum output signal, the local gradient feature descriptor is find in each position of (t,f) and calculated gradient magnitude m(t,f) and orientation θ(t,f) using by given equations

$$m(t,f) = \sqrt{d_t(t,f)^2 + d_f(t,f)^2} \quad (6)$$

$$\theta(t,f) = tan^{-1}\frac{d_t(t,f)}{d_f(t,f)} \quad (7)$$

Where $d_t(t,f) = p(t+1,f) - p(t-1,f)$, $d_f(t,f) = p(t,f+1) - p(t,f-1))$ (8)

Step 4: Finally overlapping square grids of cells are weighted the SHOG descriptors. In SHOS each square blocks are carried like building elements. The gradient magnitude and orientation are calculated the each orientation spectral histogram and quantized SHOG feature vector H. [1].

$h(\theta') = \sum_x \sum_y m(x,y)\delta[\theta', \theta(x,y)]$ (9) where θ is a set of orientation bin.

REFERENCES

[1]. Anuradha R. Fukane, Shashikant L. Sahare, "Noise estimation Algorithms for Speech Enhancement in highly non-stationary Environments", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011. ISSN (Online): 1694-0814.
[2]. Navneet Dalal and Bill Triggs, "Histogram of Oriented Gradient for Human Detection", Computer Vision and Pattern Recognition, IEEE, Computer Socienty conference, Volume-1, Page 886-893, 2005.