INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	ANJAN KUMAR TALUKDAR
Roll Number	:	10610202
Programme of Study	:	Ph.D.
Thesis Title:		Movement Epenthesis Detection in Compressed and Uncompressed Sign Language Videos
Name of Thesis Supervisor(s)	:	Prof. M.K. Bhuyan
Thesis Submitted to the Department/ Center	:	Electronics and Electrical Engineering
Date of completion of Thesis Viva-Voce Exam		:27/06/2022
Key words for description of Thesis Work	:	Continuous sign language, Movement Epenthesis, Compressed domain, Fingerspellings spottings

SHORT ABSTRACT

Continuous sign language recognition (SLR) system suffers from a problem called coarticulation. Movement epenthesis (me) is a special type of coarticulation which is a transition from one sign to other. So, sign spotting is a process of segmenting out signs and fingerspellings from continuous utterance sentences by removing movement epenthesis components, which is a challenging task.

Most of the state-of-the-art approaches on vision-based SLR systems were developed for uncompressed raw videos. However, additional information like motion vectors (MVs), macroblocks (MBs), Discrete Cosine Transform (DCT) coefficients, etc. are available in compressed domain frameworks. To address the drawbacks of the existing uncompressed domain algorithms, a framework for continuous fingerspelling spotting in the compressed domain is proposed in Chapter 3. The framework is based on motion vectors extracted from H.264/AVC compressed videos. A spatio-temporal Markov random field (ST-MRF) based model is employed to model the non-rigid motions of fingers as "sign" or "*me*". For considering camera motions, a novel framework for global motion estimation (GME) and global motion compensation (GMC) in the compressed domain is proposed. An average spotting rate of 75% is achieved.

In the second approach, a continuous fingerspelling recognition system in the uncompressed domain is proposed in Chapter 4. Adaptive thresholding and Finite State Machine (FSM) models are employed for efficient classification of "sign" and "*me*" frames. Segmented fingerspelled letters are further recognized by a conditional random field (CRF) based classifier and an average recognition rate of 91.29% is achieved.

In the third work, a framework for continuous sign language spotting using a 2-state Hidden Markov Model (HMM) with Gaussian emission probability is proposed in Chapter 5, Features extracted from the entire sign-sequence video are inputted to the HMM and finally, the hidden state-sequence is decoded using the Viterbi algorithm. From the decoded state sequence, the sign spotting is done. Experimental results show that the proposed method can separate "sign" and "me" frames with an average spotting rate of 80%.