

# Lip language identification via Wavelet entropy and K-nearest neighbor algorithm

Ran Wang<sup>1</sup>, Yifan Cui<sup>1</sup>, Xinyu Gao<sup>1</sup>, Wei Chen<sup>1</sup>, Mingbo Hu<sup>1</sup>, Qian Li<sup>1</sup>, Jiahui Wei<sup>1</sup>, XianWei Jiang<sup>1,\*</sup>

<sup>1</sup>School of Mathematics and Information Science, Nanjing Normal University of Special Education, Nanjing 210038, China

## Abstract

**INTRODUCTION:** Image processing technology is widely used in lip recognition, which can automatically detect and analyse the unstable shape of human lips.

**OBJECTIVES:** In this paper, we propose a new algorithm using Wavelet entropy (WE) and K-nearest neighbor (KNN) improves the accuracy of lip recognition.

**METHODS:** At present, the two most commonly used technologies are wavelet transform and K-nearest neighbor algorithm. Wavelet transform is a set of image descriptors, and the K-nearest neighbor algorithm has high accuracy. After a large number of experiments, we propose a lip recognition method based on Wavelet entropy and K-nearest neighbor, which combines Wavelet entropy, K-nearest neighbor and K-fold cross validation.

**RESULTS:** This method reduces the calculation time and improves the training speed. The best result of the experiment improves the accuracy to 80.08%.

**CONCLUSION:** Therefore, our algorithm is superior to other state-of-the-art approaches of lip recognition.

**Keywords:** Lip language identification, Wavelet entropy, K-nearest neighbor, Wavelet transform, K-fold cross validation

Received on 29 June 2021, accepted on 05 August 2021, published on 11 August 2021

Copyright © 2021 Ran Wang *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/\_\_\_\_\_

\*Corresponding author. Email: [jxw@njts.edu.cn](mailto:jxw@njts.edu.cn)

## 1. Introduction

### 1.1. What is Lip language identification

Lip speech recognition is a technology that combines machine vision and natural language processing to identify speech content directly from the image of someone speaking. Lip recognition system using machine vision technology, continuous identify faces from the image, determine which is the speaker, to extract the person mouth change characteristics of continuous, then enter the characteristics of continuous variation to the lip recognition model, identify the corresponding pronunciation speech

population type, then according to identify the pronunciation, calculated that the most likely of natural language statements. In the process of lip recognition, the relationship between mouth shape and pronunciation, pronunciation and text, is not the only corresponding, there are often multiple possible alternative results, need to calculate the most possible result in real time.

### 1.2. Literatures

In recent years, image processing techniques have been extensively developed for human lip recognition, which can automatically detect and analyse the unstable shape of human lips and distinguish in real time whether the user is speaking or not. Examples include audiovisual speech

















