



# Objective Hand Complexity Comparison between Two Mandarin Chinese Cued Speech Systems

Li Liu, Gang Feng, Xiaoxi Ren, Xianping Ma

## ► To cite this version:

Li Liu, Gang Feng, Xiaoxi Ren, Xianping Ma. Objective Hand Complexity Comparison between Two Mandarin Chinese Cued Speech Systems. ISCSLP 2022 -13th International Symposium on Chinese Spoken Language Processing (ISCSLP 2022), IEEE, Dec 2022, Singapore, Singapore. pp.215-219, 10.1109/ISCSLP57327.2022.10037814 . hal-03978682

**HAL Id: hal-03978682**

**<https://hal.science/hal-03978682>**

Submitted on 8 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Objective Hand Complexity Comparison between Two Mandarin Chinese Cued Speech Systems

Li Liu<sup>1</sup>, Gang Feng<sup>2</sup>, Xiaoxi Ren<sup>3</sup>, Xianping Ma<sup>4</sup>

1 Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen, China

2 Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

3 Laboratoire IADI, INSERM, université de lorraine, Nancy, France

4 School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

liuli@cuhk.edu.cn, gang.feng@gipsa-lab.grenoble-inp.fr

## Abstract

Recently, a pilot Mandarin Chinese Cued Speech (MCCS) system, called MCCS-1 was proposed with a main characteristic that each vowel is coded by only one specific hand position, without using any hand slides to code diphthongs. Indeed, hand slides are also used in some other languages of CS to code diphthongs. In order to demonstrate that the MCCS-1 system possesses real advantages over systems using hand slides, in this work, we first propose a novel MCCS-2 by introducing hand slides to code diphthongs, and a “push out move” for ending consonants [n] or [ng] of nasalized vowels. Then, we present a multi-parameter hand complexity measure method to compare MCCS-1 and MCCS-2 by measuring three kinematic parameters, which are the time duration of words realization, hand move trajectory length and average speed of hand movements. Moreover, the first MCCS corpus for these two systems is recorded by three speakers specifically for this work. A statistical analysis of this database shows a superior performance of MCCS-1 since it costs less time and takes a shorter hand move trajectory length than MCCS-2. A strong correlation is observed between the complexity of hand movement and the time duration of word realization.

**Index Terms:** Mandarin Chinese Cued Speech, hand slides, kinematic parameters, time duration, trajectory length.

## 1. Introduction

Lip reading is one of the most common communication ways for deaf people, and it helps the deaf or hearing impaired people access the spoken speech [1, 2]. However, there still exists a problem in lip reading caused by the similarity of labial shapes [3, 4] such as the ambiguity of vowel [y] and [u]. As a result, this problem makes it difficult for deaf or hearing impaired people to access the spoken speech only by traditional oral education [5].

Many methods have been proposed to overcome this problem up to now, and most of them use hand codings to provide additional information [6]. Cued Speech (CS) was invented by Dr. Cornett [7] in 1967 to make the hearing impaired people access spoken language easier. It exploits hand cues to complement the lip reading in phonetic level. In this system, the hand coding (i.e., a combination of different hand shapes and positions near the face) complements the lip reading process to enhance the speech perception. More precisely, the hand shapes are used to code consonants, while the hand positions on one side of the face or the neck are used to code vowels. Nowadays, it is estimated that CS has been adapted to over sixty languages, such as English CS [8, 9, 10, 11, 12], and French CS

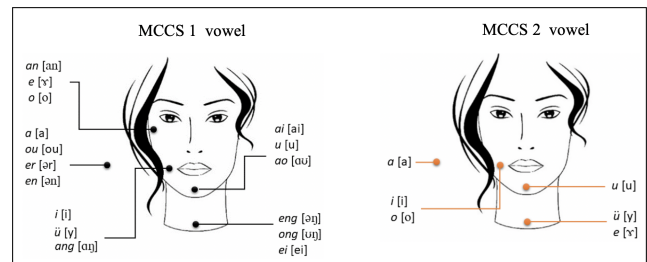


Figure 1: Comparison of vowel allocations between the proposed MCCS-2 and MCCS-1 in the previous work [16].

[13, 14, 15].

Recently, a new Mandarin Chinese CS system (MCCS) was proposed for Mandarin Chinese [16], which is called MCCS-1 in the current paper. In this system, a key approach that all the compound finals starting with *i* [i], *u* [u], *ü* [y] are coded by semi-consonants [j], [w], [ɥ] was proposed to reduce the number of vowels from 35 to 15. This is because the compound finals are coded by hand shapes instead of hand positions. Then each of remained 15 vowels is allocated to one specific hand position, without using any hand slides<sup>1</sup> to code diphthongs or compound finals. This system satisfies two main criteria [17] that: 1) hand and lips coding should be complementary (i.e., the phonemes which have similar lip shape should be distinguished by different hand shapes or hand positions); 2) cuers should spend as less as possible energy to code.

To the best of our knowledge, there are no research works concerning the CS system optimization based on the cuer’s hand kinematics analysis. However, some works are dedicated to the kinematics and kinetics of hand or arm movements in other related domains. For example, it was studied the hand movement towards a target in [18, 19, 20], and the kinematic features of the arm movements were studied in [21, 22, 23]. Particularly, in [24], Grujic and Bonkovic described measurement and analysis of human hand kinematics, and [25] analyzed kinematic parameters for handwriting. Moreover, Nelson [26] described in detail physical principles for economics of skilled movements, in which the jaw movements during speech were analyzed. Inspired by these studies, we adopted three fundamental kinematic parameters, i.e. time duration, motion trajectory length, tangential speed to the study of the CS cuer’s hand movement.

In order to demonstrate that MCCS-1 system which does not use hand slides, possess advantages over systems in which

<sup>1</sup>Hand moves from one position to another.

hand slides are used, we propose an novel alternative Mandarin Chinese CS system called MCCS-2 (see Fig. 1) using hand slides to code diphthongs. Besides, to avoid additional hand positions to code nasalized vowels, their ending [n] and [ng] are coded by two existing hand shape combined with a “push out move”. To see the performances of MCCS-1 compared with MCCS-2, under the above mentioned second criteria (the first criterion is not considered since these two systems are both satisfied), we present a multi-parameter hand complexity measure method by comparing three kinematic parameters, which are the time duration for the cuer to realize a word, the spatial trajectory length of the hand movement, and the average speed of hand movement.

## 2. A Novel Alternative Mandarin Chinese CS System

All 35 Mandarin Chinese finals and their abbreviations are listed in Table 1. Six of them are monophthongs (i.e., *a*, *o*, *e*, *i*, *u*, *ü*). Four are diphthongs (see elements in gray background), and five nasalized vowels<sup>2</sup> (see elements in blue background). Besides, there are 20 compound finals beginning with *i*, *u*, *ü* (see elements in yellow, green and orange backgrounds). Note that in this work, all monophthongs, diphthongs and nasalized vowels are referred to as vowels, while the rest are called compound finals beginning with *i*, *u*, *ü*.

MCCS-2 is designed based on three considerations: 1) By introducing hand slides, only 6 monophthongs are coded by 4 hand positions, as shown in Fig. 1 rather than 15 vowels coded by 5 hand positions in MCCS-1. 2) If all diphthongs and compound finals are coded by the hand slides, 24 hand slides would be needed, which leads to a complicated CS system. To reduce this huge amount of hand slides, the method using *i*, *u*, *ü* semi-consonant in MCCS-1 is inherited in MCCS-2, resulting in only 8 hand slides. 3) To avoid introducing extra hand positions to code nasalized vowels, we propose to code the ending consonants by a “push out” movement from two existing hand shapes, so that nasalized vowels can be realized easily and contrastively.

Table 1: 35 Mandarin Chinese finals/vowels. Simple diphthongs in gray background are called **type-1 vowels**, diphthongs beginning with *i*, *u*, *ü* in green are **type-2 finals**, simple nasal vowels in blue are **type-3 vowels**, and compound nasalized vowels beginning with *i*, *u*, *ü* in orange are **type-4 finals**.

	<i>i</i>	<i>u</i>	<i>ü</i>
<i>a</i>	<i>ia</i>	<i>ua</i>	
<i>o</i>		<i>uo</i>	
<i>e</i>	<i>ie</i>		<i>üe</i>
<i>ai</i>		<i>uai</i>	
<i>ei</i>		<i>ui(=uei)</i>	
<i>ao</i>	<i>iao</i>		
<i>ou</i>	<i>iu(=iou)</i>		
<i>an</i>	<i>ian</i>	<i>uan</i>	<i>üan</i>
<i>en</i>	<i>in(=ien)</i>	<i>un(=uen)</i>	<i>ün</i>
<i>ang</i>	<i>iang</i>	<i>uang</i>	
<i>eng</i>	<i>ing(=ieng)</i>	<i>ueng</i>	
<i>ong</i>	<i>iong</i>		

<sup>2</sup>Nasalized vowels are composed of single vowels and an ending consonant [n] or [ng].

Table 2: Composition of the corpus (40 words).

1-1	ai mei	bao tou	kao bei	shou kai
1-2	sao cui	hei niu	kuai zou	xiao mai
1-3	chai feng	long zhao	fen pei	shan kou
1-4	cai guan	pao qiang	huang mei	zhou bin
2-2	guai niao	tui xiao	xiu rui	huai jiu
2-3	diu lang	wai meng	nan qiao	dui kong
2-4	kuai dian	gui lin	xiao huang	liu run
3-3	ben teng	dong sang	san pen	song ban
3-4	qian fang	guang fan	rong jun	ceng jing
4-4	chuang ming	chuan jin	zuan nian	shun wen

## 3. Hand Complexity Measurement

To objectively describe the hand complexity in CS, we propose a hand complexity measure method by considering mainly three kinematic parameters, i.e., the word time duration, the trajectory length, and the average speed of hand movement. In the following, we firstly present our new database, and then introduce the hand complexity measure method, which is used to analyse the hand movements by the three measured kinematic parameters.

### 3.1. Database

The first MCCS corpus is specifically recorded based on MCCS-1 and MCCS-2 by three native Chinese speakers (two women and one man). CS videos are recorded in a room environment with a camera in front the CS cuer.

In the corpus, single vowels, their combinations with *i*, *u*, *ü*, and four Chinese tones are not considered since their coding methods are the same for the two systems. For the rest, according to their intrinsic features, we first divide them into 4 groups, which are type-1 vowels, type-2 finals, type-3 vowels, type-4 finals (see Table 1). We then choose Chinese characters from these 4 groups. Our corpus contains 40 words (see Table 2), each composed of two characters. In fact there are ten possible combinations (i.e., 1-1, 1-2, 1-3, 1-4, 2-2, 2-3, 2-4, 3-3, 3-4, 4-4). There are 4 words in each case, which are designed to cover all vowels and consonants with a relative good balance. Besides, to guarantee a smooth coding process, these words are inserted in a fixed sentence structure. For example, concerning the word *gui lin*, the sentence is *zhe shi gui lin ma?* Based on the above mentioned text corpus, each sentence was coded in the two MCCS systems, each containing 40 Mandarin Chinese sentences. For video recording, frame rates are 30fs and the RGB image size is 720×1280×3. In this recording, we assume that all three CS cuers have similar proficiency on these two systems to guarantee a fair comparison.

### 3.2. Measurement Method

For the time duration  $T$  of a word, the selected time interval is  $[t_1, t_2]$  and  $T = t_2 - t_1$ , where  $t_1$  is the beginning instant of the second word *shi* in each sentence, and  $t_2$  is the beginning instant of the last word *ma* in each sentence.

Concerning the hand movement trajectory length  $L$ , we first manually determine the hand back point for all images in the corpus (about 16000 images). Then we apply the spline interpolation [27] to smooth the  $x$  and  $y$  position of these hand back points. Since the distance between face and camera is variable, we use the distance between two inner eye corner  $D_{eye}$  of the

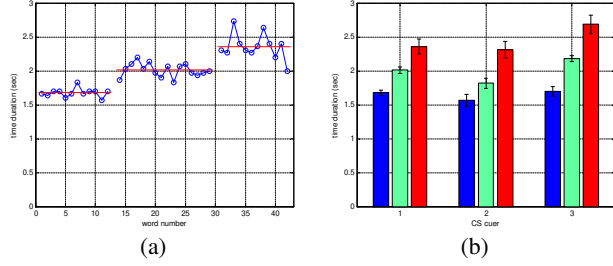


Figure 2: Time duration measure for MCCS-1 in function of the prior complexity level. (a): time duration for all 40 words (cue 1). (b): time duration and the confidence interval for all three cues. The three bars for each cue correspond to easy, middle and difficult level, respectively.

cue to normalize  $L$ :

$$L = \sum_{t=1}^{t_2-t_1} \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2} / D_{eye},$$

where  $(x_t, y_t)$  is the coordinate of the hand back position at time  $t$ . The average tangential speed  $v = \sqrt{v_x^2 + v_y^2}$ , where  $v_x$  and  $v_y$  are the speed along  $x$  and  $y$  direction, determined by the spline smoothing procedure.

### 3.3. Prior Complexity Indicator

Firstly, we present a prior knowledge on the corpus complexity. One of the main complexities for CS cue comes from the hand shape change when coding a semi-consonant i.e.,  $i, u, \ddot{u}$  between a preceded consonant and a following vowel. Therefore, if both two characters of a word do not contain the semi-consonants, this word is relatively easy to be realized. If only one of the two characters contains semi-consonants, the complexity level is middle. And if both two characters use semi-consonants, the case is considered as difficult. Following this principle, we classify all ten groups (see Table 2) of words in three complexity levels, which is shown in Table 3.

Table 3: Difficulty level of ten groups of words.

Easy	1-1, 1-3, 3-3
Middle	1-2, 1-4, 2-3, 3-4
Difficult	2-2, 2-4, 4-4

## 4. Experiment Results and Analysis

In this section, the complexity of these two systems will be examined mainly by the word time duration<sup>3</sup> and hand move trajectory length. We start with the MCCS-1 because its complexity is only dominated by the hand shape changes when the cue uses semi-consonants to code  $i, u, \ddot{u}$ , while MCCS-2 depends on other factors as well. Then, hand complexity of the MCCS-2 and MCCS-1 will be compared.

<sup>3</sup>the time duration, and spatial trajectory length in this paper means the average time duration (spatial trajectory length) of the words that belongs to the same difficulty level.

### 4.1. Analysis on MCCS-1

**Time Duration.** Firstly, the time duration of the 40 words in MCCS-1 for cue 1 is shown in Fig. 2(a) according to the prior complexity level (i.e., easy, middle and difficult). We can see a significantly intra-group homogeneity for these three groups at same complexity level. However, the time duration of different complexity level are evidently different. It is shown in Fig. 2(b) that the time duration increases successively dependent on the easy, middle and difficult complexity levels for all three cues. The differences between these three groups are larger than their 95% confidence interval (CI), indicating that the differences between groups are statistically significant. These results show that the complexity caused by the use of semi-consonants can be well measured by the time duration.

**Spatial Hand Move Trajectory Length.** On the other hand, we calculate the spatial trajectory length of hand movement. The results grouped by the prior complexity level concerning the cue 1 is shown in Fig. 3(a). It shows a stronger variations compared with the time duration distribution in Fig. 2(a). This is because the spatial length is more dependent on vowel positions than the time duration. For all three cues, the time duration and the CI of the hand movement' spatial trajectory length corresponding to three difficulty levels are shown in Fig. 3(b). We can see that globally the spatial trajectory length are almost independent on the prior complexity level, although an insignificant small variation can be observed for the cue 1. We hypothesize that all words in our database need the same number of hand position changes when coding with MCCS-1. This makes the hand movement length almost insensitive to the difficulty levels, which is directly linked to the use of semi-consonants. Fig. 3(b) shows that for different cues, the hand movement spatial trajectory length are rather different. It is natural that each cue has their personal movement characters, resulting in a specific trajectory length.

**Average Speed of Hand Movement.** Concerning the average movement speed for all three speakers, the results are shown in Fig. 4. For each cue, the results are grouped by the prior complexity levels. As the hand movement trajectory length varies few with the complexity level, while the time duration increases with the complexity level, it is natural that the average speed decreases with the complexity level. This means that for all CS cues, when the task difficulty increases, the cue takes more time to accomplish the task while decreasing their movement speed.

To sum up, the strong correlation between time duration and the prior complexity level shows that the time duration is a trustworthy measurement to indicate the difficulty of the cues. Moreover, it is more precise (weak CI) than other two measurements and the results are more homogeneous for different cues.

### 4.2. Comparison between MCCS-1 and MCCS-2

When comparing MCCS-2 with MCCS-1, two key questions arise: 1) Do the hand slides used to code diphthongs in MCCS-2 increase the hand complexity? 2) How does the ending consonants [n] and [ng] for nasalized vowels affect the hand complexity?

**Time Duration.** Time duration of the 40 words for MCCS-2 with respect to MCCS-1, for all the three cues and grouped by prior complexity level, is presented in Fig. 5. Note that the red bars corresponding to MCCS-2 are partially overlapped by the blue bars of MCCS-1, so that only the exceeded parts (i.e., time duration increment of MCCS-2 with respect to MCCS-1) are visible. Globally, time duration increases significantly com-

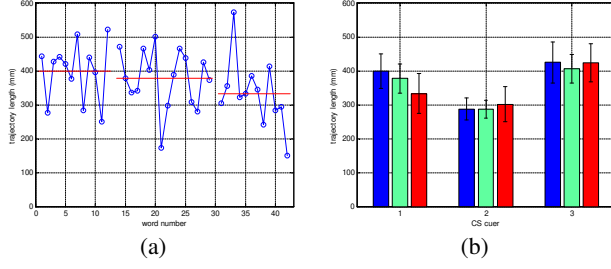


Figure 3: Trajectory length measure for MCCS-1 in function of the prior complexity level. (a): trajectory length for all 40 words (cuer 1) and (b): trajectory length and the confidence interval for all three cuers. The three bars for each cuer correspond respectively to easy, middle and difficult level.

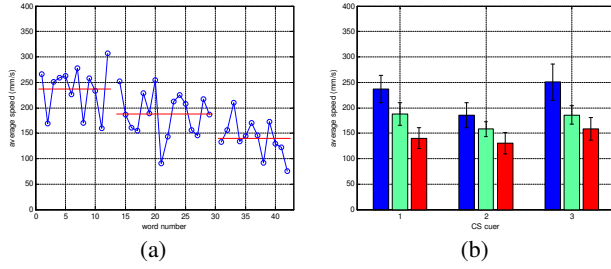


Figure 4: Average speed measure for MCCS-1 in function of the prior complexity level. (a): average speed for all 40 words (cuer 1) and (b): average speed and the confidence interval for all three cuers. The three bars for each cuer correspond respectively to easy, middle and difficult level.

pared MCCS-2 with MCCS-1. More precisely, over 32 % for cuer 1, over 64 % for cuer 2 and over 25 % for cuer 3. This means that the cuers need to spend more time to realize each word, showing a greater complexity of MCCS-2 than MCCS-1. In fact, in MCCS-2, all diphthongs are coded using a hand slide (from one position to another), and for nasalized vowels, a “push out move” to code the ending consonant [n] or [ŋ]. These additional hand movements take more time to be accomplished, directly causing an increment in the word time duration.

On the other hand, for each cuer, we observe the time duration increment from the prior complexity level 1 (easy) to level 3 (difficult) is comparable for these two MCCS systems. In details, for cuer 1, this increment is 0.68s for MCCS-1 and 0.58s for MCCS-2. For cuer 2, we can find 0.75s for MCCS-1 and 0.83s for MCCS-2. These two values are respectively 0.99s and 0.69s for cuer 3. In fact, the time duration increment from “easy” level to “difficult” level is directly caused by the use of semi-consonants coding *i*, *u*, *ii*. Since this semi-consonants rule is used for both the two systems, it appears natural that the time duration increment for both two systems is similar.

**Spatial Trajectory Length.** We calculate the spatial trajectory length of the 40 words for MCCS-2 with respect to MCCS-1 for all three cuers. Similar to MCCS-1, hand movement trajectory length is not very sensitive to prior complexity level for MCCS-2. We can observe a global increase in the trajectory length for MCCS-2 compared with MCCS-1. More precisely, 61% more for cuer 1, 73% for cuer 2 and 46% for cuer 3 (see Table 4). In fact, four target positions are needed for each word in MCCS-2 (due to the use of diphthongs and [n]/[ŋ]) rather than

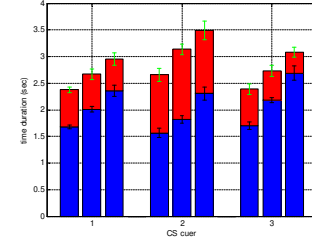


Figure 5: Time duration of all three cuers respectively for MCCS-1 (blue) and MCCS-2 (red). For MCCS-2, only the parts which exceed the blue bars are visible. The three bars for each cuer correspond to easy, middle and difficult level, respectively.

two in MCCS-1, this naturally increases the trajectory length.

**Average Speed of Hand Movement.** Given that time duration and trajectory length of MCCS-2 increase compared with MCCS-1, we are interested in how the average speed of hand movement behaves. Only the global average speed of 40 words are considered, without distinguishing prior complexity level. The ratio of the average speed for MCCS-2 over that for MCCS-1 ( $v_2/v_1$ ) is shown in Table 4 for all three cuers. We can see that compared with a relatively high ratios  $L_2/L_1$  and  $T_2/T_1$ , the ratio  $v_2/v_1$  remains low for all three cuers. This seems to show that when a cuer processes a complex system with additional hand movements, such as MCCS-2, they will spend more time to complete the task, while maintaining a relatively constant speed in hand movement. The slight variation observed in hand movement speed is coherent with the fact that all three cuers have almost the same proficiency for these two MCCS systems in our experiment.

Table 4: Comparison of the time duration ( $T$ ), trajectory length ( $L$ ) and average speed ( $v$ ) for MCCS-1 and MCCS-2. For each cuer, three ratios i.e.,  $T_2/T_1$ ,  $L_2/L_1$  and  $v_2/v_1$  are calculated, showing the increment of these parameters for MCCS-2 with respect to MCCS-1.

	cuer 1	cuer 2	cuer 3
$T_2/T_1$	1.32	1.64	1.25
$L_2/L_1$	1.61	1.73	1.46
$v_2/v_1$	1.21	1.05	1.17

## 5. Conclusion

In this work, a novel MCCS-2 using hand slides to code diphthongs, and a “push out move” to code nasalized vowels is firstly proposed. Then we propose an objective hand complexity measure method to compare MCCS-2 and MCCS-1 by three kinematic parameters, which are time duration of words, spatial trajectory length and the average speed. Evaluation on a new specially recorded MCCS corpus by three CS cuers shows that MCCS-2 costs more time for words realization (about 25% to 64% increase compared with MCCS-1), and the trajectory length of MCCS-2 is increased by 46% to 73% compared with MCCS-1, showing a higher hand complexity. Besides, the hand movement average speeds present a relatively small variation (5% to 21%). The coherence of experimental results for all three cuers shows the correctness of the proposed method to objectively compare the complexities of the two MCCS systems.



## 6. Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 62101351), the Guangdong Basic and Applied Basic Research Foundation (No.2020A1515110376), Shenzhen Outstanding Scientific and Technological Innovation Talents PhD Startup Project (No. RCBS20210609104447108).

## 7. References

- [1] B. E. Dodd and R. E. Campbell, *Hearing by eye: The psychology of lip-reading*. Lawrence Erlbaum Associates, Inc, 1987.
- [2] M. F. Woodward and C. G. Barber, "Phoneme perception in lipreading," *Journal of Speech, Language, and Hearing Research*, vol. 3, no. 3, pp. 212–222, 1960.
- [3] T. Ezzat and T. Poggio, "Miketalk: A talking facial display based on morphing visemes," in *Proceedings Computer Animation'98 (Cat. No. 98EX169)*. IEEE, 1998, pp. 96–102.
- [4] L. Liu, G. Feng, and D. Beutemps, "Automatic dynamic template tracking of inner lips based on clnf," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 5130–5134.
- [5] L. Liu, "Modeling for continuous cued speech recognition in french using advanced machine learning methods," Ph.D. dissertation, Université Grenoble Alpes, 2018.
- [6] N. Pugeault and R. Bowden, "Spelling it out: Real-time asl fingerspelling recognition," in *2011 IEEE International conference on computer vision workshops (ICCV workshops)*. IEEE, 2011, pp. 1114–1119.
- [7] R. O. Cornett, "Cued speech," *American annals of the deaf*, vol. 112, no. 1, pp. 3–13, 1967.
- [8] L. Liu, G. Feng, D. Beutemps, and X.-P. Zhang, "A novel resynchronization procedure for hand-lips fusion applied to continuous french cued speech recognition," in *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 2019, pp. 1–5.
- [9] —, "Re-synchronization using the hand preceding model for multi-modal fusion in automatic continuous cued speech recognition," *IEEE Transactions on Multimedia*, vol. 23, pp. 292–305, 2020.
- [10] L. Liu, J. Li, G. Feng, and X.-P. Zhang, "Automatic detection of the temporal segmentation of hand movements in british english cued speech," *Proc. Interspeech 2019*, pp. 2285–2289, 2019.
- [11] J. Wang, N. Gu, M. Yu, X. Li, Q. Fang, and L. Liu, "An attention self-supervised contrastive learning based three-stage model for hand shape feature representation in cued speech," *Proc. Interspeech 2021*, pp. 626–630.
- [12] J. Wang, Z. Tang, X. Li, M. Yu, Q. Fang, and L. Liu, "Cross-modal knowledge distillation method for automatic cued speech recognition," *Proc. Interspeech 2021*, pp. 2886–2990.
- [13] L. Liu, G. Feng, and D. Beutemps, "Extraction automatique de contour de lèvres à partir du modèle clnf," in *Actes des 31èmes Journées d'Etude de la Parole*, 2016.
- [14] —, "Inner lips feature extraction based on clnf with hybrid dynamic template for cued speech," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, pp. 1–15, 2017.
- [15] —, "Visual recognition of continuous cued speech using a tandem cnn-hmm approach," in *11th Annual Conference of the International Speech Communication Association (Interspeech)*, 2018.
- [16] L. Li and F. Gang, "A pilot study on mandarin chinese cued speech," *American Annals of the Deaf*, vol. 164, no. 4, pp. 496–518, 2019.
- [17] R. O. Cornett, "Adapting cued speech to additional languages," *Cued Speech Journal*, vol. 5, pp. 19–29, 1994.
- [18] W. Beggs and C. Howarth, "The movement of the hand towards a target," *Quarterly Journal of Experimental Psychology*, vol. 24, no. 4, pp. 448–453, 1972.
- [19] P. J. Bairstow, "Analysis of hand movement to moving targets," *Human Movement Science*, vol. 6, no. 3, pp. 205–231, 1987.
- [20] J. Messier and J. F. Kalaska, "Comparison of variability of initial kinematics and endpoints of reaching movements," *Experimental Brain Research*, vol. 125, no. 2, pp. 139–152, 1999.
- [21] C. G. Atkeson and J. M. Hollerbach, "Kinematic features of unrestrained vertical arm movements," *Journal of Neuroscience*, vol. 5, no. 9, pp. 2318–2330, 1985.
- [22] X. Xiao, H. Hu, L. Li, and L. Li, "Comparison of dominant hand to non-dominant hand in conduction of reaching task from 3d kinematic data: Trade-off between successful rate and movement efficiency," vol. 16, no. 3, pp. 1611–1624, 2019.
- [23] H. N. Zelaznik, R. A. Schmidt, and S. C. Gielen, "Kinematic properties of rapid aimed hand movements," *Journal of Motor Behavior*, vol. 18, no. 4, pp. 353–372, 1986.
- [24] T. Grujić and M. Bonković, "Measurement and analysis of human hand kinematics," *World Academy of Science, Engineering and Technology International Journal of Biomedical and Biological Engineering*, vol. 9, no. 2, p. 97, 2015.
- [25] A. Accardo, M. Genna, and M. Borean, "Analysis of handwriting kinematic parameters in italian children," in *V Latin American Congress on Biomedical Engineering CLAIB 2011 May 16-21, 2011, Habana, Cuba*. Springer, 2013, pp. 1114–1117.
- [26] W. L. Nelson, "Physical principles for economies of skilled movements," *Biological cybernetics*, vol. 46, no. 2, pp. 135–147, 1983.
- [27] I. J. Schoenberg, *Cardinal spline interpolation*. Siam, 1973, vol. 12.