International Journal of Advanced Intelligence
Volume 3, Number 2, pp.269-285, July, 2011

A New Framework to Deal with OOV Words in SLT System

Yu Zhou, Feifei Zhai ,Peng Liu, Chengqing Zong

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences

No. 95, Zhongguancun East Road, Room 1010, Zidonghua Building, Beijing 100190, China

{*yzhou, ffzhai, pliu,cqzong}@nlpr.ia.ac.cn*}

Received (10-12-2010)
Revised (15-3-2011)

Automatic spoken language translation (SLT) is considered as one of the most challenging tasks in modern computer science and technology. It is always a hard nut to deal with the problem of Out-Of-Vocabulary (OOV) words in SLT. The existing traditional SLT framework often doesn't take effect for OOV word translation because of the data sparseness. In this paper based on the analysis of common OOV expressions appeared in SLT, we propose a new framework for bidirectional Chinese-English SLT in which a series of approaches to translating OOV expressions are presented. The experimental results have shown that our framework and approaches are effective and can greatly improve the translation performance.

Keywords: Spoken language translation; OOV; named entity translation; digital and time named entity.

1. Introduction

The purpose of spoken language translation (SLT) is to make a computer system work like a human interpreter for two different language speakers. SLT is considered as one of the most challenging tasks in modern computer science and technology [Kitano, 1994; Zong et al., 2005].

As we know a typical SLT system consists of three main key modules: Automatic Speech Recognizer (ASR), Machine Translator (MT), and Text-To-Speech synthesizer (TTS). The three modules are typically integrated in a pipeline structure that is shown in Figure 1.

Compared with text-to-text translation, SLT has the following unique characteristics: ① the average length of spoken language sentences is much shorter than text sentences and the structure is relatively simpler. According to our statistics, the average length of spoken Chinese sentences is about 7.8 Chinese words and the average word length in spoken Chinese is about 1.87 characters, but in Chinese text the average length of sentences is about 22 Chinese words and each word contains about 2.5 Chinese characters in average [Zong et al., 1999].

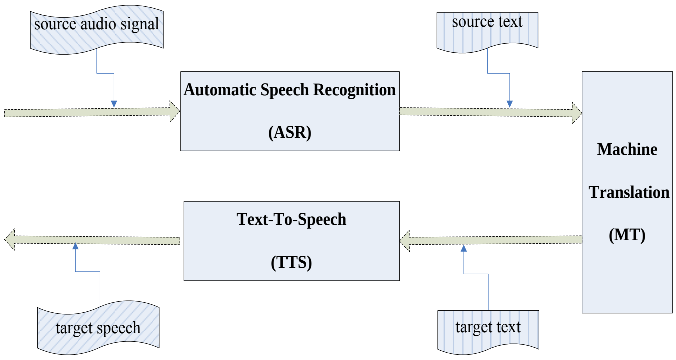


Fig. 1. A Typical Architecture of SLT System

It is quite different; ② there are so many ill-formed expressions in spoken language utterances, including redundancy, repetition, repair, word order confusion, ellipsis, and so on. According to our statistics, in spoken Chinese language, about 4.70% sentences contain redundancy, 3.56% sentences have repetition, 32.61% sentences are incomplete, and 44.59% sentences are only of one word [Zong et al., 1999]; ③ In Chinese-to-English translation, a large number of homonyms in Chinese characters or Chinese words make it impossible to have a very satisfactory recognition results. Therefore, it is a very challenging task to correctly translate the results with noise from ASR; ④ a practical SLT system needs real time response speed. However, it is difficult to have a high speed in such an integrated system with several modules.

In recent years, advances in ASR, MT, and TTS technologies have paved the way for the emergence of various SLT systems and ensured the basic performance and robustness. However, though SLT research and development have made significant progresses over the decades, there still exist a number of theoretical and technical problems and still has much room to improve the system performance.

This paper presents a new approach to dealing with the problem of OOV words in Chinese-to-English SLT. The remainder of the paper is organized as follows: Section 2 gives our investigation of OOV words in Chinese-to-English SLT; Section 3 introduces the related work and Section 4 gives our framework by introducing our approach to dealing with OOV in bidirectional Chinese-English SLT; Section 5 gives the details on implementation of our approaches; The experimental results are shown in Section 6. And finally, the concluding remarks are given in Section 7.

2. Investigation of OOV Words in SLT

Due to the characteristics of spoken language, there are always numbers of OOV words in SLT. According to our investigation, the OOV words in Chinese-English translation engine using statistical machine translation (SMT) models are mainly caused by the following reasons: a) The limited size of training data: the training data is always insufficient and the sparseness is always a serious problem in SMT; b) For Chinese-to-English SLT, due to the problem of homonyms in Chinese, it is a common type of ASR errors that a Chinese word or a character is wrongly substituted by its homonyms. And also, different speaker with different accent or pronunciation style, the errors are quite different. For example, if a Chinese speaker with southern accent in mainland of China utters the following sentence "冰冻三尺，非一日之寒。(the Chinese Pinyin is: bing dong san chi, fei yi ri zhi han)"(Rome was not built in a day.), the ASR result is probably "冰冻三次，非一日自汗", in which the Chinese words underlined are wrongly given because they have the similar pronunciation with their original words "三尺" and "之寒" respectively; c) There are often some new words in a new area. Even if in the same area, there will be new words with the change of time. For example, "给力(feed force), 奥特(out of date/out of fashion)" are all new words in spoken Chinese in recent years. Therefore, it is impossible to make the training corpus and the extracted translation rules cover all of the language phenomena and lexical information.

We have analyzed 461 OOV words in 2,652 Chinese ASR output sentences. The types of OOV words are shown in Table 1.

Table 1. Statistics on OOV Words in ASR Output

OOV Types	Number	Ratio (%)
Personal names	178	38.61
Place names	39	8.46
Organization names	22	4.77
Digits	69	14.97
Date&Time	10	2.17
Foreign language words	73	15.84
Others	70	15.18
Total	461	100.00

Having further investigated the 461 OOV words, we found that most of the OOV words are caused by the wrong word boundaries. And the wrong word boundaries are caused by the wrong syllable segmentation. Table 2 gives the statistical results on distribution of incorrect boundaries.

Table 2. Statistics on Incorrect Boundaries of OOV Words

OOV Type	Number	Boundary Error	Ratio(%)
Personal names	178	14	7.87%
Place names	39	4	10.26%
Organization names	22	11	50.00%
Digits	69	2	2.90%
Time&Date	10	5	50.00%
Others	70	9	12.86%
Total	388	45	-

From the statistical results in Table 1 and Table 2, we can clearly see that the OOV words are widespread in ASR outputs. As shown in Table 1 and Table 2, the OOV problem is mainly caused by the following reasons: (1) Named entities (NE) including the personal names, place names and organization names, which we call such OOV as OOV-1; (2) Digits, date and time OOVs, we call such OOV expressions as OOV-2; (3) Homophone words which mainly generated by the wrong ASR output, we call such OOV words as OOV-3; (4) Other OOV words, which are signed as OOV-4; and (5) Foreign language words, which are named as OOV-5.

3. Related Work

Many approaches have been proposed to solve the problem of OOV words, in which the interactive translation approach is more popular. The research on interactive translation has been carried out for many years since [Kay et al., 1973] first implemented the interactive translation system MIND, such as [Waibel, 1996; Seligman, 1997; Zong et al., 2002, 2005] and so on. According to the different stages of interaction that user is involved in, we divide the interactive approaches into three types as follows: ① Interaction before translation (IBT). The basic principle of IBT is to correct the wrong or ambiguous recognition output of ASR before translation and make the MT input correct and unambiguous. Generally, the tasks of IBT method include to format source text, segment source sentences, correct the recognition errors, process the special symbols, and so on [Waibel, 1996; Seligman, 1997]. ② Interaction during translation (IDT). IDT method carries out interaction in translation process and generates candidate translations [Lane et al., 2008; Zong et al., 2002, 2005]. and ③ Interaction after translation (IAT). IAT method is the simplest one. Its main objective is to provide users with a friendly interface to let users choose proper translation results from candidates or correct the translation errors without much effort. The users to use this method have to know both the source language and target language. The representative systems are SYSTRAN* and HICATS/JE [Kaji, 1987]. Three girls in Figure 2 stand for the different interaction methods in SLT process respectively.

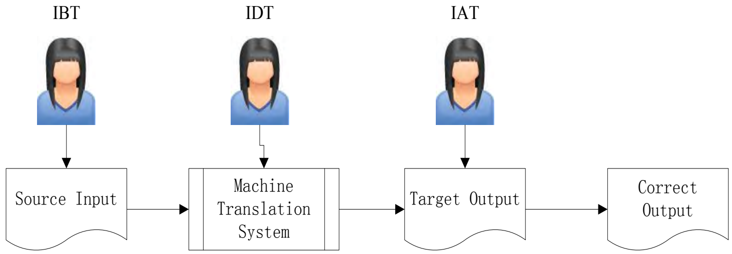


Fig. 2. The Three Interaction Types

In some approaches the additional language resources are employed to solve the OOV problem. [Zhou et al., 2007] proposed an approach to interpreting the semantic meanings of OOV words by using the synonyms knowledge in the source language. This method is only effective to alleviate those OOV words that have the synonyms with the same Part-Of-Speech (POS). If the synonyms have different POSs, the translation engine may get wrong translation results.

[Habash, 2008] proposed the following four methods to deal with the problem of OOV words: (a) Spelling expansion; (b) Morphological expansion; (c) Dictionary term expansion; and (d) Proper name transliteration. However, most of the methods are not suitable for bidirectional Chinese-English translation system for Chinese is not an adhesive language without inflection.

[Mirkin et al., 2009] proposed a method to use WordNet to solve the problem of OOV words. [Aziz et al., 2010] proposed an approach to taking into account of both paraphrased and entailed words and used a context model score.

All of these related work mentioned above have their own merits. This brings up the question: how to make good use of the merits of existing approaches to improve the performance of a Chinese-English SLT system? As we mentioned above, IBT method and IAT method are aimed to processing the input and output of translation module respectively. They are not involved in the process of machine translation. As we said before, even if an input is completely correct, the MT engine probably still can't translate it due to the ill-formed expressions or other problems [Zong et al., 2005]. IDT method works in MT process, but it never pre-processes the input even if an input is completely wrong or not understandable. So we propose a new framework

---

* http://www.systransoft.com/

to solve the problem of OOV words by combing the three interaction methods, which makes full use of the merits of the three interaction methods.

4. Our Framework

Our frameworks to deal with the problems of OOV words for Chinese-to-English (C2E) translation and English-to-Chinese (E2C) translation are respectively shown in Figure 3 and Figure 4. From the figures we can see that our approaches to dealing with OOV problem are different from the existing SLT framework shown in Figure 1. In Figure 3, in each stage there are three types of output (word-based, character-based, and Pinyin-based) which can greatly alleviate the OOVs. In the final stage, a combination modular is employed to find an optimal translation as output given to TTS modular.

Based on the frameworks given in Figure 3 and Figure 4, the SLT systems work as the following procedure:

(i) A source speech is recognized into the source text by an ASR;
(ii) All NEs in source text are identified and translated first, then the original NEs are replaced with their corresponding translations (target NEs);
(iii) The digits, date and time expressions in source language are translated into target ones by D&T recognition and translation module;
(iv) The new generated source text is translated into target text by MT module after step (2) and (3);
(v) There may still exist some OOV words without translation, they will be replaced by using common bilingual translation dictionaries;
(vi) Finally, the target text is converted into speech by TTS module.

In order to recognize and translate the OOV-1, OOV-2 and OOV-4 more accurately, we simultaneously introduce three modes (word-based, character-based and Pinyin-based) in IBT/IAT stages on CE SLT system. How can the three modes alleviate the OOV problems effectively? It will be fully reflected in the later translation and combination model, shown in Section 5. In IDT stage, we propose a new translation model to alleviate the OOV problems and such model is especially very effective for the OOV-3. The new translation model is built based on three word alignments, which are word-based, Pinyin-based, and character-based respectively. And it will play a more important role by introducing a combination model in IAT stage. The combination model is imperative to generate an optimal translation result based on the candidate translations from different translation engines. In Section 5, we will focus on C2E translation to give more details. In summary, our approaches to dealing with the different types of OOV words are given in detail as follows:

OOV-1 Processing: In SLT system, especially in text-to-text SMT system, NEs recognition and translation are usually processed by an individual module using some special approaches, e.g., the work proposed by [Chen et al., 2008]. But if we introduce such method in SLT system and translate each NE depending on the method, it will heavily increase the complexity of the overall system and reduce the whole translation speed. Even when we introduce a NE recognition and translation module, such problem still exists. So here we introduce the IBT idea to automatically recognize and translation the NEs of ASR output in advance. Our method works in the following two steps: 1) We only translate the source NEs with a maximum-matching algorithm with support of a bilingual NE database. Here the bilingual NE database also provides three types information. The three types are word-based, character-based and Pinyin-based. 2) We use the method proposed by [Chen et al., 2008] to translate the complex NEs, for example, the organization names.

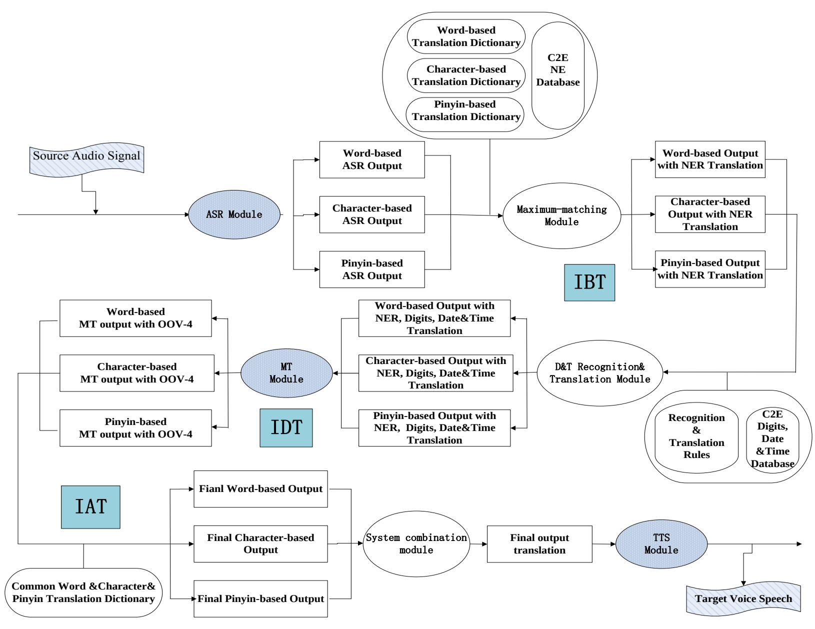


Fig. 3. Our Framework to Deal with OOV Words in C2E SLT System

OOV-2 Processing: For digits, date and time OOV expression recognition and translation, we have developed a special tool named as D&T module [Zhai et al., 2009]. Zhai et al. (2009) carefully investigated the structural characteristics of time and number named entities in both Chinese and English and classified them into several kinds. They have developed a rule-based analyzer to recognize and translate all digits, date and time expressions from Chinese to English and English to Chinese.

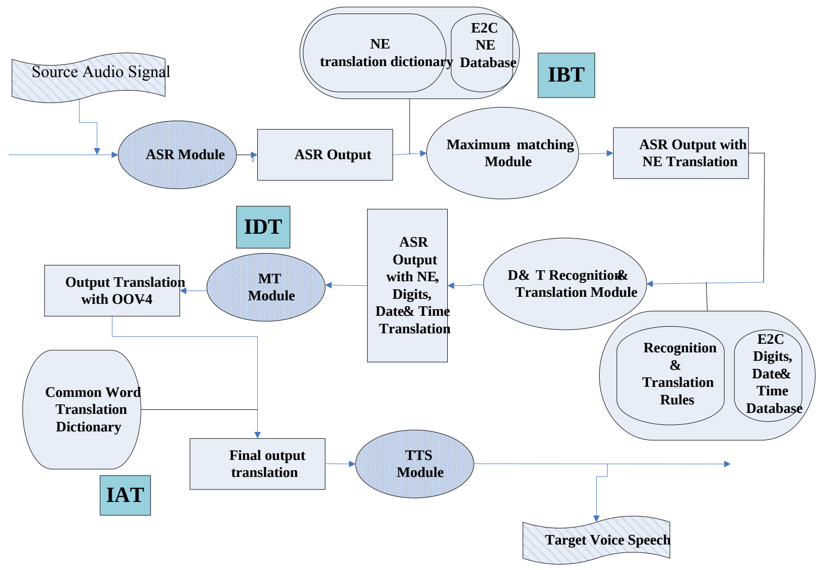


Fig. 4. Our Framework to Deal with OOVs in E2C SLT System

OOV-3 Processing: This type of OOV words is processed by an additional module using Pinyin-based module to alleviate such problem. See Section 5.

OOV-4 Processing: This type of OOV words is processed by using a common bilingual dictionary.

OOV-5 Processing: This type of OOV words is directly output as the final translations.

5. Implementation of Our Approaches

To implement a complete MT system, the key techniques mainly include two issues: one is to extract the translation knowledge (including the translation rules and re-ordering rules) in ***training process*** and the other one is to develop an effective decoding algorithm in ***decoding process***. At present, the rule-based method and statistical method are two basic methods and the statistical method has become the mainstream in recent years. In statistical method there are word-based, phrase-based, and syntax-based translation models as well. The common nature of these models is to extract more complete and accurate translation knowledge from training data so as to guide the decoder to get the better translation results. Now the decoding algorithms mainly use certain search strategy, which guides the decoder to find an optimal path to obtain the final translation results with the guidance of pre-extracted translation knowledge. Considering the different translation engines are probably based on different translation models and get different translation results, some combination approaches have been proposed to emerge the translation results from different translation engines to generate the better translation result as final output. The process of system combination is called as ***combining process*** hereafter.

Now we give the corresponding framework of the three main technologies embodied in the processes of training, decoding and combining, which are shown in Figure 5 a), b) and c) respectively. The operating steps of each process will be described separately. The shadow modules in these figures are the highlighted objects which are modified by our proposed methods. These methods will be described later in detail.

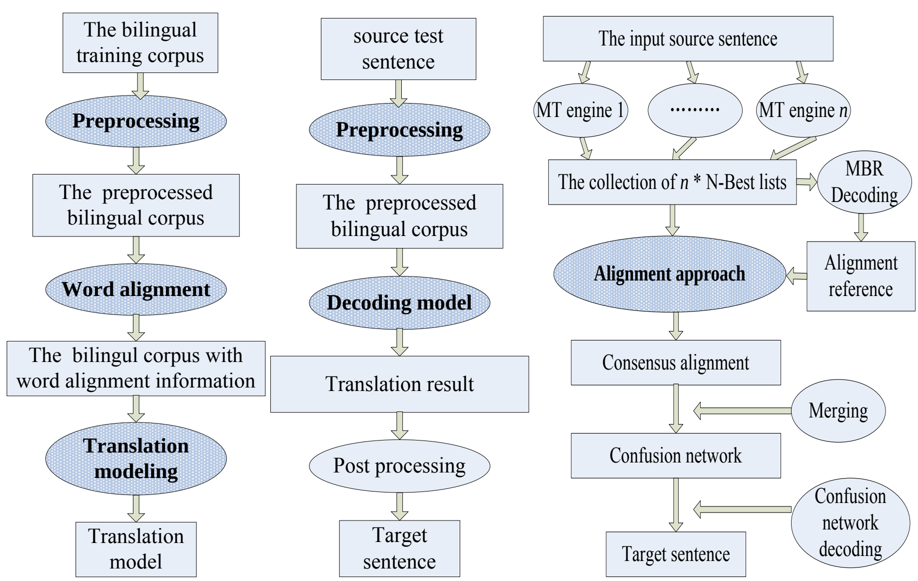


Fig. 5. a). Training process Figure 5 b). Decoding process Figure 5 c). Combining process

- **Training process**

As shown in Figure 5 a), the training process works through the following steps:

(i) The bilingual training corpus is pre-processed including the Chinese word segmentation, English tokenization, and so forth;

(ii) The pre-processed bilingual training data are given to the word alignment model to obtain the word alignments;
(iii) The bilingual training data with the word alignments information are passed to the translation model to extract the translation knowledge as the translation model.
(iv) The language model is generally obtained by the free toolkit (such as SRILM†) without any modification.

- **Decoding process**

As shown in Figure 5 b), the decoding process works through the following steps:
(i) The test sentence is pre-processed including the Chinese word segmentation, English tokenization, and so forth;
(ii) The pre-processed test sentences are sent to the decoding model to obtain the translation result;
(iii) The translation result is then post-processed to be final translation.

- **Combining process**

As shown in Figure 5 c), the combining process works through the following steps:
(i) Input sentence is pre-processed and passed into multiple MT engines (suppose n MT engines) to produce n*N-Best lists of translations separately from n MT engines;
(ii) The alignment reference is obtained by the MBR Decoding;
(iii) The consensus alignment is built with the help of alignment approach and the alignment reference;
(iv) The confusion network is built based on the consensus alignment;
(v) The target sentence is generated by using the confusion network decoding.

Currently, there are few methods proposed to process the noisy results from ASR module. In order to better adapt to our framework, we have proposed efficient methods and solutions, which are highlighted with shadow shown in Figure 5. The modified methods are described as follows.

5.1. *Training &Decoding Process*

Traditionally, the original bilingual data are expressed in one type, such as word-based or character-based. In order to solve the OOV-3 problem effectively, we introduce the type of Pinyin-based input. In order to obtain more accurate and complete translation knowledge, we pre-process the ASR output into three types, namely, word-based, character-based, and Pinyin-based. This idea is shown in Figure 6. In the following, we will first explain why we can improve the translation performance by introducing the three types of translation models. Then we will explain why we can solve OOV-3 problem with Pinyin-based model.

---

† http://www.speech.sri.com/projects/srilm/manpages/

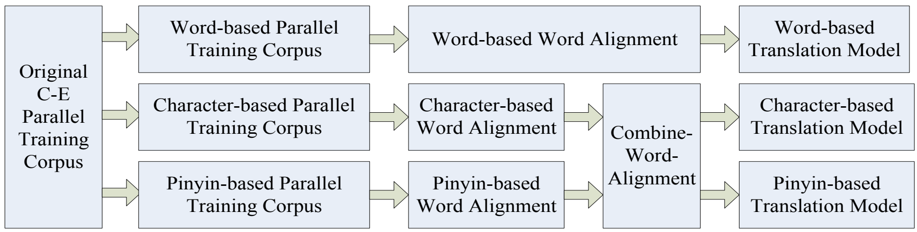


Fig. 6. MT Module Based on Three Types of Inputs

Based on our experience on MT development, a different word alignment can extract different translation knowledge and combining different word alignments will lead to achieve better translation knowledge. We here combine the character-based and Pinyin-based word alignments for the character and the Pinyin have the same positions in a sentence. As we all know, we can achieve a better translation result by combining more translation hypotheses. So here we use the three types of input instead of one to obtain the corresponding three translation models and to generate more translation hypotheses.

The current translation models are based on literal form (word-based or character-based model). They cannot recall the type of OOV-3 which is caused by ASR errors. Now we use an example to show why we can solve OOV-3 problem by introducing Pinyin-based translation model for C2E translation, see Figure 7.

From the example in Figure 7, we can clearly see that it can also get correct translation results even with a wrong recognition output by ASR. The reasons are summarized as the following two points from both training and decoding viewpoints:

我	的	名	字	是	**铃**	木	直	子

a) The correct speech recognition results

我	的	名	字	是	**玲**	木	直	子

b) The error speech recognition results

wo3	de5	ming2	zi4	shi4	**ling2**	mu4	zhi2	zi5

c) the pinyin output of a) and b)

my	name	is	naoko	suzuki

d) the corresponding correct translation results

Fig. 7. Comparison on Chinese Character-based and Pinyin-based Translation Models

(i) ***In training stage***, suppose there are two parallel sentences. Taking {a),d)} and { b),d)} in Figure 7 as examples, we can find that the literal form of "铃木直子" is different from "玲木直子". In word alignment of training stage, it may receive lower probabilities aligned to the corresponding target translation "*naoko suzuki*". More precisely, if there exist more times of "铃木直子" but few times of "玲木直子", the probability of "玲木直子" linked to "*naoko suzuki*" will be too low to extract such translation pair. But if we use Pinyin "*ling2 mu4 zhi2 zi5*" instead of the literal form of both "铃木直子" and "玲木直子", the translation probability of "*ling2 mu4 zhi2 zi5*" to "*naoko suzuki*" will be higher. Thus we recall such translation pairs.

(ii) ***In decoding stage***, if a sentence contains "玲木直子", it cannot find the matching accurate target translation which will greatly decrease the translation quality. But if we add Pinyin module, we can transfer the "玲木直子" to "*ling2 mu4 zhi2 zi5*" to find the accurate translation.

5.2. *Combining Process*

Now we have three translation models: word-based model, character-based model and Pinyin-based model and *3*N-Best* candidate translation lists will be generated. In order to better play the role of the three translation models, it is natural that we will depend on the combination model. There are alternative approaches for system combination based on word level and sentence level. Here we choose the word level approach because this approach can achieve comparatively a better result. In the word level approach, the most important sub-module is the word alignment. Here the word re-ordering alignment (WRA) is used to address the word alignment with different valid word orders. The WRA approach can directly shift the word sequences of the translation hypothesis to the correct location within the translation hypothesis [Li et al., 2008].

6. Experiment

For translation engine, we evaluate the performance on both translation speed and quality. First, we give the statistics of the bilingual training data, which are used to train the language model and translation model. Those bilingual training data are collected from the travel guidebook, daily communication handbook, business communication handbook, and internet etc. The detailed statistics are shown in Table 3. The test data use 919 sentences of ASR result and the corresponding correct text sentences. For each source test sentence, we only give one target reference translation linked with the source test sentence.

Table 3. Statistics of Bilingual Training Data

Language	Sentences	Number of words	Number of characters
Chinese	1,100,197	31,416	5,225
English	1,100,197	27,739	---

In order to compare the influence of recognition accuracy, we give a comparison of translation results between the inputs of speech recognition results and correct text respectively. Table 4 gives the results. Here BLEU [Papineni et al., 2002] is the most popular measure metric in evaluation of machine translation.

From Table 4 we can see that from the comparison of ASR results and the correct text, the ASR results have a great influence on the quality of translation. For the same input correct text, the BLEU score is decreased about 10 points by the ASR errors.

Table 4. Comparison on Different Inputs

Language	Input		BLEU	Translation time (s/Sen.)	Av. Len. of Sen. (Word)
C2E	ASR	male	0.5316	0.16645	6.9880
		female	0.5184	0.165992	7.0392
	TXT		0.5944	0.17385	5.8945
E2C	ASR	male	0.3678	0.179866	7.0968
		female	0.3715	0.180733	7.1523
	TXT		0.4805	0.21025	6.3308

We make a comparison on the test data of IWSLT2008 [‡] (International Workshop on Spoken Language Translation, 2008). The purpose of this comparison is to further explain the importance of ASR accuracy and the sensitivity of training data. The comparison results are shown in Table 5. From Table 5, we can clearly see that the different test data cause big difference in BLEU score with the same translation model and language model. It shows that the translation performance (BLEU score) is significantly decreased when the training data and test data are more dispersive.

We also give the comparison results by introducing Pinyin-based model for the C2E translation. Our experimental data are mainly from IWSLT2009 [§] and the comparison is done under the free toolkits Moses [**] with default settings. Table 6 gives the data statistics on training, development, and test data. Table 7 gives the

‡ http://mastarpj.nict.go.jp/IWSLT2008/
§ http://mastarpj.nict.go.jp/IWSLT2009/
** http://www.statmt.org./moses/

comparison of translation results based on three modes: word-based, Pinyin-based and character-based model. Table 7 also gives the combination results based on WRA. From Table 7, we can see that translation performance is greatly improved by introducing Pinyin-based translation model and combination model. By the Pinyin-based translation model, the BLEU score has been improved for about 3 points compared to the word-based translation model for it greatly alleviates the OOV-3 problem which is caused by the ASR module. By introducing the combination model, the BLEU score has been greatly improved for the combination model can choose an optimal path to generate a better translation and can reduce the OOV error by a voting method based on many translation candidates generated by three translation models.

Table 5. Comparison of MT Results with Different Inputs

Inputs	BLEU	Translation time (s/Sen.)	Av. Len. of Sen. (Number of words)
Chinese text	0.2839	0.158257	7.445
Chinese ASR results	0.2493	0.17586	6.520
English text	0.4215	0.143557	7.281
English ASR results	0.3289	0.110024	5.542

Table 6. Data Statistics Used in Training, Development, and Test Sets

Corpus data	Training data (Pair)	Development data	Test data
Sentences	30,033	4,447	405

Table 7. Comparison of Translation Results Based on Word or Pinyin

Translation model	BLEU score on development data	BLEU score on test data
Word-based	33.48	29.65
Pinyin-based	36.43	32.04
Character-based	33.78	30.12
Combination (WRA)	38.22	34.31

7. Conclusions

In this paper we propose a new framework to deal with the problem of OOV words in SLT, in which the same output from ASR is expressed and pre-processed by character-based, word-based, and Pinyin-based approach. This framework has the following strong points: (1) it takes advantage of the merits of three interactive types and can exert different advantages in the various stages; (2) it can easily import a variety of external resources in each translation stage; (3) in IAT stage, it can alleviate the OOV-1 and OOV-2 problem by introducing the NE and D&T recognition and translation module; (4) it can greatly alleviate the OOV-3 problem by introducing the Pinyin-based model; (5) it combines three types of word alignments which can greatly improve the accuracy of translation model; and (6) it can find an optimal translation result by introducing the combination model. The experimental results have shown that our framework can greatly improve the translation performance and have good robustness.

Acknowledgments

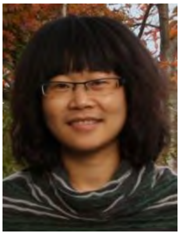
The research work has been partially funded by the Natural Science Foundation of China under Grant No. 6097 5053 and 61003160, and also supported by the External Cooperation Program of the Chinese Academy of Sciences as well.

References

1. Aziz, Wilker, Marc Dymetman, Shachar Mirkin, Lucia Specia, Nicola Cancedda, and Ido Dagan. 2010. Learning an Expert from Human Annotations in Statistical Machine Translation: the Case of Out-of-Vocabulary Words. In Proceedings of EAMT, Saint-Raphael, France.
2. Chen,Yufeng, Chengqing Zong. 2008. A Structure-based Model for Chinese Organization Name Translation. ACM Transactions on Asian Language Information Processing, Mach 2008, 7(1): 1-30.
3. Habash, Nizar. 2008. Four Techniques for Online Handling of Out-of-Vocabulary Words in Arabic-English Statistical Machine Translation, In *Proceedings of ACL-08: HLT, Short Papers*, pp. 57-60.
4. Kaji, Hiroyuki. 1987. HICATS/JE: a Japanese-to-English machine translation system based on semantics. In *MT Summit*, Japan, pp. 55-60.
5. Kay, Martin. 1973. The MIND System. Natural Language Processing.
6. Kitano, Hiroaki: 1994, Speech-to-speech Translation: A Massively Parallel Memory-Based Approach,Kluwer Academic Publishers, Boston.

7. Lane, Ian R. / Waibel, Alex (2008): "Class-based statistical machine translation for field maintainable speech-to-speech translation", In INTERSPEECH-2008, 2362-2365.
8. Li, Maoxi and Chengqing Zong. Word Re-ordering Alignment for Combination of Statistical Machine Translation Systems. In Proceedings of the International Symposium on Chinese Spoken Language Processing (ISCSLP), December 16-19, 2008. Kunming, China.
9. Mirkin, Shachar, Lucia Specia, Nicola Cancedda, Ido Dagan, Marc Dymetman, and Idan Szpektor. 2009. Sourcelanguage entailment modeling for translating unknown terms. In Proceedings of ACL, Singapore.
10. Papineni, K., Roukos, S., Ward, T., and Zhu, W. J. (2002). "BLEU: a method for automatic evaluation of machine translation" in *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics* pp. 311–318
11. Seligman, Mark. 1997. Interactive real-time translation via the Internet. In: Working notes, natural language. In Proceedings of AAAI-97 Spring Symposium on Processing for the World Wide Web, Stanford, CA, pp 142–148
12. Waibel, Alex. 1996. Interactive translation of conversational speech. *Computer*, 29(7):41–48
13. Zhai, Feifei, Rui Xia, Yu Zhou, Chengqing Zong. 2009. An Approach to Recognizing and Translating Chinese & English Time and Number Name Entities (in Chinese). In *Proceedings of the fifth China Workshop on Machine Translation*, pp.172-179. October 16-17, 2009, Nanjing, China
14. Zhou, Keyan, Chengqing Zong. 2007. Dealing with OOV Words in Chinese-to-Englsih Statistical Machine Translation System (in Chinese). In *Proceedings of the 9th Chinese National Conference on Computational Linguistics*. Dalian, August 6-8, 2007. Pages 356-361.
15. Zong, Chengqing, Hua Wu, Taiyi Huang, and Bo Xu. 1999. Analysis on Characteristics of Chinese Spoken Language. In Proceedings of 5th Natural Language Processing Pacific Rim Symposium (NLPRS). November 1999, Beijing. Pages 358-362
16. Zong, Chengqing, Bo Xu, and Taiyi Huang. 2002. Interactive Chinese-to-English speech translation based on dialogue management. In Proceedings of ACL 2002 Workshop on Speech-to-Speech Translation: Algorithms and Systems, Philadelphia, Pennsylvania, pp 61–68
17. Zong, Chengqing and Mark Seligman. 2005. Toward Practical Spoken Language Translation. *Machine Translation*, 19(2): 113-137.

First Author



Yu Zhou was graduated in 1999 from Northeastern University, China and received her M.S. degree from her Alma Mater in 2002. In 2008, she got her Ph. D degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences (CASIA). After she got Ph. D degree, she joined NLPR and now works as an engineer. Her current research interests include machine translation and the key techniques of natural language processing.

Second Author

Feifei Zhai was graduated in 2009 from Beijing Jiaotong University, China and then became a Ph. D candidate of the Institute of Automation, Chinese Academy of Sciences (CASIA). His current research direction is machine translation.

Third Author

Peng Liu was graduated in 2002 from Xi'an Jiaotong University, China and received his M.S. degree from his Alma Mater in 2005. In 2010, he got his Ph. D degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences (CASIA). His current research interests include machine translation and the key techniques of natural language processing.

Fourth Author

Chengqing Zong is a professor in natural-language technology at the National Laboratory of Pattern Recognition and the deputy director of the National Laboratory of Pattern Recognition, which is part of the Chinese Academy of Sciences' Institute of Automation. He's also a guest professor at the Graduate University of the Chinese Academy of Sciences. His research interests include machine translation, sentiment classification, and human-computer dialogue systems. Zong received his Ph. D from the Chinese Academy of Sciences' Institute of Computing Technology in March 1998. He's a director of the Chinese Association of Artificial Intelligence and the Society of Chinese Information Processing, and is an executive member of the Asian Federation of Natural Language Processing.