

# 神经机器翻译系统在维吾尔语-汉语翻译中的性能对比

哈里旦木·阿布都克里木, 刘洋, 孙茂松

(清华大学 计算机科学与技术系, 智能技术与系统国家重点实验室, 清华信息科学与技术国家实验室(筹), 北京 100084)

**摘要:** 基于深度学习的神经机器翻译已在多个语言对上显著超过传统的统计机器翻译, 成为当前的主流机器翻译技术。该文从词粒度层面出发, 对国际上具有影响力的6种神经机器翻译方法在维吾尔语-汉语翻译任务上进行了深入分析和比较, 这6种方法分别是基于注意力机制(Ground-Hog), 词表扩大(LV-groundhog), 源语言和目标语言采用子词(subword-nmt)、字符与词混合(nmt.hybrid)、子词与字符(dl4mt-cdec)以及完全字符(dl4mt-c2c)方法。实验结果表明: 源语言采用子词、目标语言采用字符的方法(dl4mt-cdec)在维吾尔语-汉语神经机器翻译任务上性能最佳。该文不仅是首次将神经机器翻译方法应用到维吾尔语-汉语机器翻译任务上, 也是首次将不同的神经机器翻译方法在同一语料库上进行了对比分析。该研究对维吾尔语-汉语机器翻译任务和神经机器翻译的进一步研究工作都具有重要的参考意义。

**关键词:** 神经机器翻译; 资源匮乏语言; 维吾尔语

中图分类号: TP391.2

文献标志码: A

文章编号: 1000-0054(2017)08-0878-06

DOI: 10.16511/j.cnki.qhdxxb.2017.22.054

## Performance comparison of neural machine translation systems in Uyghur-Chinese translation

Halidanmu Abudukelimu, LIU Yang, SUN Maosong

(Tsinghua National Laboratory for Information Science and Technology,

State Key Laboratory of Intelligent Technology and Systems,

Department of Computer Science and Technology,

Tsinghua University, Beijing 100084, China)

**Abstract:** The neural machine translation based on deep learning significantly surpasses the traditional statistical machine translation in many languages, and becomes the current mainstream machine translation technology. This paper compares six influential neural machine translation methods from the level of word granularity in the task of Uyghur-Chinese machine translation. These methods are attention mechanism (GroundHog), vocabulary expansion (LV-groundhog), source language and target language with subword units (subword-nmt), characters and words mixed (nmt.hybrid), subword units and characters (dl4mt-cdec), and complete characters (dl4mt-c2c). The experimental results show that Uyghur-Chinese neural machine

translation performs best when the source language is segmented into subword units and the target language is represented by characters (dl4mt-cdec). This paper is the first to use neural machine translation for Uyghur-Chinese machine translation and the first to compare different neural machine translation methods on the same corpus. This work is an important reference not only for Uyghur-Chinese machine translation, but also for general neural machine translation tasks.

**Key words:** neural machine translation; low-resource language; Uyghur

神经机器翻译(neural machine translation, NMT)是近几年兴起的机器翻译方法<sup>[1-2]</sup>。神经机器翻译的核心思想是使用一个循环神经网络将源语言句子编码为一个稠密向量, 然后从该向量解码出目标语言句子。由于神经机器翻译能够通过长短期记忆和注意力等机制有效处理长距离依赖问题, 其翻译性能在汉语-英语、英语-法语、英语-德语、俄语-英语和捷克语-英语等多个语言对上显著超过了传统的统计机器翻译(statistical machine translation, SMT)。因此, 神经机器翻译迅速取代统计机器翻译成为当前主流的机器翻译技术<sup>[3-4]</sup>。

目前, 国际上提出了多种神经机器翻译方法。文[2]提出了基于注意力机制(attention)的神经机器翻译方法。文[5]提出词表扩大的方法。文[6]采用了未登录词替换的方法。文[7-11]提出基于最小意义单元表示建模。其中, 最具代表性的是文[12]提出的字节对编码(byte pair encoding, BPE)技术。文[13-15]提出了多任务、多语言、zero-shot等方法。虽然已有研究针对低资源、形态丰富语言的神经机器翻译提出了多种解决方法, 但是这

收稿日期: 2017-02-23

基金项目: 国家自然科学基金重点项目(61331013);

国家“八六三”高技术项目(2015AA015407)

作者简介: 哈里旦木·阿布都克里木(1978—), 女, 博士研究生。

通信作者: 刘洋, 副教授, E-mail: liuyang2011@tsinghua.edu.cn

些研究着眼的角度不同,差异很大,也未在公开的同—数据集上进行深入分析与对比,因此无法知道这些方法在不同语言环境下的有效性和局限性,这使得无从综合评价哪一种方法更适合维吾尔语-汉语机器翻译任务。

维吾尔语作为一种黏着语在语素的组合上具有高度的灵活性,虽然词干和词缀的数量有限,但是理论上可以组合生成无限多的词语,绝大多数维吾尔语词语在语料库中只出现一次<sup>[16-17]</sup>。维吾尔语通过在词干上添加词缀来实现丰富的句法和语义功能。以表 1 为例,“باغچىلاشتۇرۇماقچىمىز”(即“我们准备建园林”)作为一个词汇却表示出通常意义下句子层面的信息,这种情况在维吾尔语-汉语机器翻译中造成了严重的数据稀疏问题。

表 1 维吾尔语语言特性举例

维吾尔语词	汉语译文
باغ	园
باغچا	花园
باغچىلاش	花园化
باغچىلاشتۇر	使花园化
باغچىلاشتۇرماق	建园林
باغچىلاشتۇرماقچى	准备建园林
باغچىلاشتۇرۇماقچىمىز	我们准备建园林

更为严重的是,与维吾尔语相关的平行语料库严重匮乏。目前,英语、汉语等资源丰富语言的双语语料规模已达亿级以上,但是可公开购买使用的维吾尔语-汉语平行语料库仅有 5 万句对的小规模。由于神经机器翻译模型参数规模远远大于统计机器翻译模型,如果没有充足的训练数据,其翻译质量甚至会远远低于统计机器翻译。

本文将 6 种神经机器翻译方法在维吾尔语-汉语机器翻译任务上进行了全面细致的比较,深入分析了这些方法在处理不同语言现象时的有效性和局限性。实验结果表明,在维吾尔语-汉语机器翻译任务中,神经机器翻译对维吾尔语不需要事先进行任何形态切分也能获得很好的翻译效果。编码器和解码器都使用字符表示的方法<sup>[11]</sup>严重依赖于语言特性。编码器使用 BPE、解码器使用字符的方法<sup>[9]</sup>在维吾尔语-汉语机器翻译任务上效果最好。本文工作首次在同一语料库上对不同的神经机器翻译方法进行了对比,这对进一步开展神经机器翻译研究工作具有重要的参考意义。

## 2 6 种有影响力的神经机器翻译系统简介

近年来,国际学术界提出了多种神经机器翻译

方法,并发布了相应的开源系统,最具有代表性的是 Bahdanau 等<sup>[2]</sup>提出的基于注意力机制的神经机器翻译方法。该方法拉开了神经机器翻译超越统计机器翻译的大幕,其核心思想是在循环神经网络(recurrent neural network, RNN)框架下引入了注意力机制,表示源语言词语和目标语言词语的关联强度。对应的系统为 GroundHog。

随着目标词表的扩大,神经机器翻译的训练和解码的复杂度增加,对于大规模的语料,使用全部的目标词很困难。于是, Jean 等<sup>[5]</sup>使用采样算法(sampling)扩大了训练词表,进而提升了机器翻译质量,对应的系统为 LV-groundhog。

但是, Jean 等<sup>[5]</sup>的方法对于在训练语料中未出现的词仍不能识别。Sennrich 等采用了 BPE 技术<sup>[12]</sup>,提出了子词方法。其核心思想是语料库中的高频词作为整体出现的可能性大,因此不对其进行切分,而对低频词进行切分,由此增加稀疏词中子词的共现次数。该方法有效解决了数据稀疏问题,对应的系统为 subword-nmt。BPE 技术已在工业界的机器翻译技术中使用<sup>[3]</sup>。Sennrich 等采用该技术在 WMT2016 机器翻译评测的 8 个语言对中的 7 个上获得了第 1 名。本文的实验结果也发现 BPE 技术大大提高了维吾尔语-汉语机器翻译的质量。

除此之外, Luong 等<sup>[10]</sup>提出了字词混合(hybrid)方法<sup>[10]</sup>,该方法在长短时记忆(long short-term memory, LSTM)网络框架下对高频词用词表示,而对稀疏词或未登录词用字符表示。该方法的优点是既能够保留更多词层面的语义信息,同时又解决了数据稀疏问题,对应的系统为 nmt.hybrid。

以上的方法分别从词表扩大、子词以及字词混合的角度解决数据稀疏问题,但是如果能够采用基于字符的机器翻译将是更理想的解决方法。因为基于字符的方法有更多优势:1)一种语言中的任何词,无论它多么生僻,都是由该语言的字符构成的,有了该语言所有字符向量,遇到该语言的任何词都可以采用字符向量的组合来表示该词;2)采用字符表示可以大大减少空间复杂度,一般情况下,一种语言的字符总和在 200 个左右,比起万级、几十万级的词表,大大节省了存储和解码空间。Chung 等<sup>[7]</sup>和 Lee 等<sup>[11]</sup>从字符层面上提出了基于字符的神经机器翻译方法。其中, Chung 等<sup>[7]</sup>提出的方法为在 RNN 框架下源语言采用 BPE 技术编码、目标语言采用字符解码,解码端采用双规模(bi-scale)循环神经网络的模型架构。Lee 等<sup>[11]</sup>的

方法是在 Chung 等<sup>[7]</sup>的系统框架下在编码器端加入了卷积神经网络(convolutional neural network, CNN)。这两种方法对应的系统分别为 dl4mt-cdec 和 dl4mt-c2c。

以上 6 种神经翻译方法的系统框架、词粒度等信息见表 5。本文将在实验部分对这些系统的一些特性及整体性能进行评估。

### 3 实验语料与实验设置

#### 3.1 语料

本文使用的语料是新疆大学的新疆多语种信息技术重点实验室公开的维吾尔语-汉语综合领域平行语料库(<http://www.chineseldc.org/>)。该语料库包含 5 万维吾尔语-汉语句对。使用该语料主要

出于 2 个原因: 1) 语料库已发布; 2) 语料为维吾尔语-汉语句对。

#### 3.2 预处理

语料库包含 5 万维吾尔语-汉语句对, 本文采用随机抽取算法, 将语料分成了训练集(48 000 句对)、开发集(1 000 句对)和测试集(1 000 句对)。维吾尔语训练集未进行形态切分之前的词语数量为 1 166 118, 词汇量是 74 644。本文采用 Morfessor 软件(<http://morfessor.readthedocs.io/en/latest/general.html>)<sup>[18]</sup>对维吾尔语进行形态切分, 同时使用了清华大学 THULAC 软件(<http://thulac.thunlp.org/>)对汉语分词。对维吾尔语和汉语数据进行形态切分和分词后的详细信息如表 2 所示。

表 2 训练集、开发集和测试集语料库统计

数据集	语言	分词工具	句子数	词语数量	词汇量
训练集	维吾尔	Morfessor	48 000	1 834 117	10 420
	汉	THULAC	48 000	1 109 198	43 717
开发集	维吾尔	Morfessor	1 000	38 424	4 200
	汉	THULAC	1 000	23 462	5 570
测试集	维吾尔	Morfessor	1 000	38 568	4 315
	汉	THULAC	1 000	23 576	5 554

#### 3.3 实验设置

神经机器翻译实验对所有的数据进行了乱序(shuffle)操作。本文使用 BLEU(bilingual evaluation understudy)作为评测指标(<https://github.com/moses-smt/mosesdecoder/blob/master/scripts/generic/multi-bleu.perl>)。由于不同系统采用不同的架构, 很难达到一致性, 因此本文默认每个系统超参数的值是最优状态下的值。dl4mt-c2c 系统中有 2 种模型设置——双语言对和多语言对模型, 本文采用双语言对模型。所有神经机器翻译训练的迭代次数限制在 20 万次。子词词表大小均设定为 3 万。实验中, 统计机器翻译实验系统采用开源的 Moses(<http://www.statmt.org/moses/>)<sup>[19]</sup>, 该系统是基于短语(phrase-based)的统计机器翻译系统。

### 4 实验结果

#### 4.1 BPE 与形态切分的关系

为了验证使用 BPE 前是否需要维吾尔语数据进行形态切分, 本实验分别比较了在未进行形态切分的维吾尔语数据上使用 BPE 和在用 Morfessor

形态切分后的数据上再使用 BPE 这两种情况在 GroundHog 系统上的性能体现。实验结果显示, 两种数据在开发集上的 BLEU 值分别为 13.72 和 13.93, 后者比前者仅仅提高了 0.21, 提升并不显著。由此可知, 对维吾尔语使用 BPE 技术时, 可以不对维吾尔语进行形态切分, 其原因可能是 BPE 技术可以较好地解决数据稀疏问题。但由于汉语句子里中字与字之间没有空格, 如果不分词而直接使用 BPE 技术, 效果可能不理想, 因此本文没有对目标端汉语数据进行类似的实验。

#### 4.2 词表大小与形态切分的关系

本实验分析了对维吾尔语数据未进行形态切分和进行形态切分两种情况下, 在不同词表设定下的神经机器翻译方法的性能。

表 3 显示不同词粒度和词表大小设置下 GroundHog 系统的机器翻译结果。表头下方第 1 行为对维吾尔语数据进行形态切分、词表取 1 万的实验结果, 第 2 行为对维吾尔语数据进行形态切分后取全部词表(大小为 10 420)的实验结果, 第 3 行是维吾尔语数据未进行形态切分、词表大小取 3 万的实验结果。实验结果表明, 当维吾尔语数据进行

形态切分后,用部分词表比用全部词表在开发集和测试集上的 BLEU 值高一些。分析其原因可能是形态切分后的语素是离散排列的,如果将形态切分后的全部语素都包含到词表中,可能会由于低频离散语素丢失了词层面上的语义信息,从而影响机器翻译的性能。表 3 的实验结果表明,维吾尔语数据不使用形态切分、词表大小使用 3 万比使用形态切分、用部分词表或全部词表数据的 BLEU 值在测试集上分别高出 0.49 和 0.67,由此说明在低资源、形态丰富的维吾尔语-汉语数据集上,相比形态切分,词表更能提高机器翻译的性能,其原因可能是形态切分会导致词层面上的语义信息丢失。与之相反,如果扩大词表,既能减少未登录词,也保留了词层面上的语义信息,神经网络可以从中学习词的形态特征及结构,进而提高机器翻译的性能。另外,形态切分工具是独立于机器翻译的,因此切分结果的好坏直接影响机器翻译的性能。

表 3 词表大小与形态切分对 GroundHog 系统性能的影响

源语言	词表	BLEU	
		开发集	测试集
语素	10 000	13.70	12.84
语素	10 420	13.55	12.66
词	30 000	13.74	13.33

### 4.3 基于字符的 dl4mt-c2c 系统与汉语分词关系

文[11]提出了源和目标语言不需要进行任何

分词的方法,对应的系统为 dl4mt-c2c。为了验证它在汉语上的性能,本实验对汉语分词与否两种情况进行对比。

由表 4 可知,不对汉语进行分词,dl4mt-c2c 系统的效果很差,说明文[11]提出的方法并不是语言独立的,对汉语这种字与字之间没有任何空格的孤立语并不适合。文[20]指出机器翻译的评测方法 BLEU 并不适合字符级别的 NMT 系统的评估。

表 4 dl4mt-c2c 系统在中文分词与否情况下的性能

源语言	目标语言	BLEU	
		开发集	测试集
词	未分词	6.15	3.88
词	词	16.45	15.05

### 4.4 不同机器翻译方法在维吾尔语-汉语数据上的性能

本实验分析了维吾尔语-汉语数据分别在 SMT 和 NMT 上的性能体现。其中,SMT 系统采用开源的 Moses 系统<sup>[19]</sup>,对应的实验在其默认设置下完成。NMT 采用第 2 部分所述的 6 种系统。在 NMT 实验中维吾尔语不进行形态切分,汉语进行分词,GroundHog 词表大小取 3 万,子词词表大小也取 3 万,LV-groundhog 词表设置为维吾尔语-汉语的全部词表。字符表取每种语言训练集中的全部字符。实验结果如表 5 所示。

表 5 统计机器翻译及不同神经机器翻译系统的对比

系统类别	文献	系统	框架	源语言	目标语言	BLEU	
						开发集	测试集
SMT	Koehn 等 <sup>[19]</sup>	Moses	phrase-based	语素	词	28.85	27.18
	Bahdanau 等 <sup>[2]</sup>	GroundHog	RNN+attention	词	词	13.74	13.33
NMT	Jean 等 <sup>[5]</sup>	LV-groundhog	RNN+attention+sampling	词	词	13.57	13.10
	Luong 等 <sup>[10]</sup>	nmt, hybrid	LSTM+hybrid	词+字符	词+字符	13.59	13.15
	Sennrich 等 <sup>[12]</sup>	subword-nmt	BPE	子词	子词	13.72	13.57
	Lee 等 <sup>[11]</sup>	dl4mt-c2c	CNN+RNN+bi-scale	字符	字符	16.45	15.05
	Chung 等 <sup>[7]</sup>	dl4mt-cdec	RNN+BPE+bi-scale	子词	字符	19.34	18.00

由表 5 可知,目前的 NMT 系统在维吾尔语-汉语翻译上还不能超过传统的 SMT 系统,这和文[21]得出的结论一致,即 NMT 系统在豪萨语、土耳其语、乌兹别克语和乌尔都语这 4 种资源匮乏语言的小规模数据集上的性能无法超过 SMT 系统。

NMT 机器翻译系统中效果最好的是 dl4mt-cdec,其 BLEU 值比 GroundHog 高出了 4.67。为

了排除语言依赖性,本文作者在百万规模的汉语-英语数据上也进行了相同的实验,实验结果同样也是 dl4mt-cdec 的 BLEU 值最高。可见,在这 6 个 NMT 系统中,dl4mt-cdec 性能最佳,而且具有良好的语言无关性。

LV-groundhog 没有超过 GroundHog 的原因可能是本研究的维吾尔语-汉语语料规模太小,采

样技术没有很好地发挥作用。对于 dl4mt-c2c 系统, 已有研究<sup>[11]</sup> 在双语言对模型上分别对德语-英语、捷克语-英语、芬兰语-英语和俄语-英语做了实验, 只有德语-英语的 BLEU 值超过了 dl4mt-cdec。因此, dl4mt-c2c 在维吾尔语-汉语机器翻译任务上没有超过 dl4mt-cdec 是合理的, 这也说明 dl4m-c2c 系统的性能可能具有语言依赖性。nmt.hybrid 系统包含 2 个子系统, 分别对应文[22] 和文[10]。其中文[22] 是带有全局上下文注意力机制的 NMT 系统, 文[10] 在文[22] 基础上, 对于未登录词使用基于字符的双层循环神经网络架构, 是为了解决数据稀疏问题。但是, 本文作者发现文[22] 和 [10] 在维吾尔语-汉语机器翻译任务上的 BLEU 值相同, 其原因可能是文[10] 对未登录词采用基于字符的网络架构在小规模维吾尔语-汉语数据集上不能很好地发挥作用, 维吾尔语和汉语的语言差异大也可能是一个原因, 两个相近语言的翻译效果可能会好一些。

## 5 结 论

本文采用维吾尔语-汉语机器翻译任务对国际上近年来提出的 6 种具有影响力的神经机器翻译方法(基于注意力机制(GroundHog), 词表扩大(LV-groundhog), 源语言和目标语言采用子词(sub-word-nmt)、字符与词混合(nmt.hybrid)、子词与字符(dl4mt-cdec)以及完全字符(dl4mt-c2c))进行了全面深入对比。本文发现, 在维吾尔语-汉语神经机器翻译任务上, 不需要事先对维吾尔语进行形态切分, 只需扩大词表也能得到相当或更好的翻译结果。同时, 本文发现在这 6 种神经机器翻译方法中 Chung 等<sup>[7]</sup> 提出的 dl4mt-cdec 系统在维吾尔语-汉语机器翻译任务上取得了最佳效果。目前提出的完全字符的神经机器翻译方法<sup>[11]</sup> dl4mt-c2c 并不适合解码端采用孤立语如汉语的机器翻译任务, 使用该方法时仍然需要对汉语进行分词才能取得良好的翻译效果。未来, 将进一步研究低资源、形态丰富语言的神经机器翻译技术, 以提升翻译效果。

## 参考文献 (References)

- [1] Sutskever I, Vinyals O, Le Q. Sequence to Sequence Learning with Neural Networks [Z/OL]. (2014-09-10)[2015-03-15]. <https://arxiv.org/abs/1409.3215>.
- [2] Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate [Z/OL]. (2014-09-01)[2015-03-20]. <https://arxiv.org/abs/1409.0473>.
- [3] Wu Y, Schuster M, Chen Z, et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation [Z/OL]. (2016-09-26)[2016-10-02]. <https://arxiv.org/abs/1609.08144>.
- [4] Johnson M, Schuster M, Le Q, et al. Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation [Z/OL]. (2016-11-14)[2016-11-16]. <https://arxiv.org/abs/1611.04558>.
- [5] Jean S, Cho K, Memisevic R, et al. On using very large target vocabulary for neural machine translation [C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. Beijing, 2015: 1-10.
- [6] Luong M, Sutskever I, Le Q, et al. Addressing the rare word problem in neural machine translation [C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. Beijing, 2015: 11-19.
- [7] Chung J, Cho K, Bengio Y. A character-level decoder without explicit segmentation for neural machine translation [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1693-1703.
- [8] Ling W, Trancoso I, Dyer C, et al. Character-Based Neural Machine Translation [Z/OL]. (2015-11-14)[2016-02-03]. <https://arxiv.org/abs/1511.04586>.
- [9] Costa-Jussà M, Fonollosa J. Character-based neural machine translation [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 357-361.
- [10] Luong M, Manning C. Achieving open vocabulary neural machine translation with hybrid word-character models [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1054-1063.
- [11] Lee J, Cho K, Hofmann T. Fully Character-Level Neural Machine Translation without Explicit Segmentation [Z/OL]. (2016-10-10)[2016-10-22]. <https://arxiv.org/abs/1610.03017>.
- [12] Sennrich R, Haddow B, Birch A. Neural machine translation of rare words with subword units [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1715-1725.
- [13] Luong M, Le Q, Sutskever I, et al. Multi-Task Sequence to Sequence Learning [Z/OL]. (2015-11-19)[2016-07-17]. <https://arxiv.org/abs/1511.06114>.
- [14] Firat O, Cho K, Bengio Y. Multi-way, multilingual neural machine translation with a shared attention mechanism [C]// Proceedings of NAACL-HLT. San Diego, CA, USA, 2016: 866-875.
- [15] Firat O, Sankaran B, Al-Onaizan Y, et al. Zero-resource translation with multi-lingual neural machine translation [C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, TX, USA, 2016: 268-277.
- [16] 哈里旦木·阿布都克里木, 程勇, 刘洋, 等. 基于双向门限递归单元神经网络的维吾尔语形态切分 [J]. 清华大学学报(自然科学版), 2017, 57(1): 1-6. ABUDUKELIMU Halidanmu, CHENG Yong, LIU Yang, et al. Uyghur morphological segmentation with bidirectional GRU neural networks [J]. *J Tsinghua Univ (Sci and Tech)*, 2017, 57(1): 1-6. (in Chinese)

- [17] Abudukelimu H, Liu Y, Chen X, et al. Learning distributed representations of Uyghur words and morphemes [C]// Proceedings of CCL/NLP-NABD. Guangzhou, 2015: 202 - 211.
- [18] Creutz M, Lagus K. Unsupervised discovery of morphemes [C]// Morphological and Phonological Learning: Proceedings of the 6th Workshop of the ACL Special Interest Group in Computational Phonology (SIGPHON). Stroudsburg, PA, USA, 2002: 21 - 30.
- [19] Koehn P, Hoang H, Birch A, et al. Moses: Open source toolkit for statistical machine translation [C]// Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions. Pargue, Czech Republic, 2007: 177 - 180.
- [20] Sennrich R. How Grammatical Is Character-Level Neural Machine Translation Assessing MT Quality with Contrastive Translation Pairs [Z/OL]. (2016-12-14)[2016-12-20]. <https://arxiv.org/abs/1612.04629>.
- [21] Zohp B, Yuret D, May J, et al. Transfer learning for low-resource neural machine translation [C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, TX, 2016: 1568 - 1575.
- [22] Luong M, Pham H, Manning C. Effective approaches to attention-based neural machine translation [C]// Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal, 2015: 1412 - 1421.

---

(上接第 877 页)

- [12] Ikhlef A, Louveaux J. Per subchannel equalization for MIMO FBMC/OQAM systems [C]// Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim). Victoria, British, 2009: 559 - 564.
- [13] Waldhauser D S, Baltar L G, Nossek J A. Adaptive equalization for filter bank based multicarrier systems [C]// Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS). Seattle, WA, USA, 2008: 3098 - 3101.
- [14] Biglieri E. Coding and modulation for a horrible channel [J]. *IEEE Communications Magazine*, 2003, **41**(5): 92 - 98.
- [15] Nguyen H H, Bui T Q. Bit-interleaved coded OFDM with iterative decoding in impulsive noise [J]. *IEEE Transactions on Power Delivery*, 2008, **23**(2): 640 - 649.
- [16] Benedetto J J, Heil C, Walnut D F. Gabor systems and the Balian-Low theorem [M]// Gabor Analysis and Algorithms. Cambridge, MA, USA: Birkhäuser Boston, 1998: 85 - 122.
- [17] Siohan P, Siclet C, Lacaille N. Analysis and design of OFDM/OQAM systems based on filterbank theory [J]. *IEEE Transactions on Signal Processing*, 2002, **50**(5): 1170 - 1183.
- [18] Morelli M, Moretti M. Improved decoding of BICM-OFDM transmissions plagued by narrowband interference [J]. *IEEE Transactions on Wireless Communications*, 2011, **10**(1): 20 - 26.
- [19] Viholainen A, Ihalainen T, Stitz T H, et al. Prototype filter design for filter bank based multicarrier transmission [C]// Proceedings of 17th European Signal Processing Conference (EUSIPCO). Glasgow, UK, 2009: 1359 - 1363.