

融合敏感词规则和字符级 RCNN 模型的用户意图识别

王冲 张虎 王鑫 刘全明

(山西大学计算机与信息技术学院 山西 太原 030006)

摘要 用户意图识别是基于用户对话用语识别用户的真实对话意图,是人机对话研究中的一项关键任务。针对现有用户意图识别方法的不足,提出融合敏感词规则和字符级 RCNN 模型的用户意图识别方法。构建敏感句子与敏感词词典,并通过规则及相似度匹配策略对特征明显的对话进行意图识别。针对类别特征不明显的对话提出深层语义分类模型,该模型以单字符串作为输入序列,利用 RCNN 模型构建意图分类框架,既可以避免分词结果不准确带来的错误传导问题,同时利用字符的分布向量表示方法还可以获取句子的深层语义信息。实验结果表明,该方法在两个数据集上都取得了较好的结果,明显优于传统的意图识别方法。

关键词 人机对话 用户意图识别 自然语言处理 RCNN

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2020.03.027

USER INTENT RECOGNITION BASED ON SENSITIVE WORD RULES AND CHARACTER LEVEL RCNN MODEL

Wang Chong Zhang Hu Wang Xin Liu Quanming

(School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China)

Abstract User intention recognition is a key task in the research of human-machine dialogue based on real conversation intention of identifying users with dialogue language. Aiming at the shortcomings of existing methods of existing user intention recognition methods, this paper proposes a user intention recognition method, which integrates sensitive word rules and character-level RCNN model. The sensitive sentence and sensitive word dictionary were constructed, and intention recognition of dialogues with obvious features was identified by rules and similarity matching strategy. Then, a deep semantic classification model was proposed for the dialogue with inconspicuous category characteristics. The model used a single string as the input sequence, and used the RCNN model to construct the intent classification framework, which could avoid the error conduction problem caused by inaccurate word segmentation results. At the same time, the distribution vector representation method of the character could also be used to obtain the deep semantic information of the sentence. Experimental results show that the proposed method achieves good results on both data sets and is obviously superior to the traditional method of intent recognition.

Keywords Human-machine dialogue User intent recognition Natural language processing RCNN

0 引言

近年来,随着自然语言理解和深度学习技术的快速发展和广泛应用,人机对话已成为自然语言处理领域的一项重要研究任务,已引起学术界和企业界的广泛关注。目前国内外相关学术会议例如 SIGIR2018、

ICTIR2017、EMNLP2018 和 SMP2018 等都开展了人机对话研究的分组研讨或公开评测,专门就人机对话技术和应用展开了深入交流,这些会议促进了人机对话技术的发展,推动了语音识别、信息检索和自动问答等技术在人机对话系统的深入应用。当前很多公司也专门开展了人机对话的应用研究,将人机对话技术应用到企业客服、个人事务助理、虚拟情感陪护机器人以及

娱乐性聊天机器人等产品,典型的产品有 Amazon Echo、Google Home、Microsoft Cortana、百度小度、阿里的天猫精灵等。

经典的人机对话系统一般包括三个主要部分:用户意图识别、用户对话管理和自然语言生成。其中用户意图识别是人机交互系统能否进行准确、有效对话的一个首要基础环节^[1]。然而,由于人机对话中对话环境的复杂性及用户输入方式的多性样,对话系统经常对用户意图识别出错,大大影响了交互的体验。因此迫切需要针对人机对话中的用户意图识别任务提出更有效的解决办法。

目前意图识别的典型方法是将意图识别任务转化为意图的自动分类任务。即在给定意图类别体系的情况下,通过提取特征、训练分类模型来实现对意图的自动识别^[2]。例如用户输入“你好啊,很高兴见到你!”,系统将其分到闲聊类;输入“我想订一张去北京的机票”,系统将其分到“订机票”类;对于“我想找一家五道口附近便宜干净的快捷酒店”,将其分为“订酒店”类。

国外的研究人员早期曾提出使用句子相似度进行用户意图的识别^[3]。国内研究人员将特征扩展思想引入意图识别研究中,即引入额外的信息挖掘文本所表达的含义来辅助分类^[4]。这些研究都在特定数据集上取得了有效的结果,然而由于这些意图识别方法大都基于模板匹配或人工特征集合,不仅费时费力、扩展性差^[5],还经常由于意图识别错误出现答非所问的情形。

针对上述方法的不足,本文考虑了人机对话文本的长度短、噪声大和容易出现分词错误等特点,构建了融合敏感词规则和字符级 RCNN 的双层意图识别模型。首先基于部分对话文本类别词明显的特点构建了敏感句子及敏感词词典,并通过规则及相似度匹配策略对特征明显的对话进行了意图识别;然后针对类别特征不明显的对话提出了深层语义分类模型,该模型以单字符串作为输入序列,利用 RCNN 模型构建了意图分类框架,既可以避免分词结果不准确带来的错误传导问题,同时利用字符的分布向量表示方法还可以获取句子的深层语义信息。

1 相关工作

1.1 意图识别

意图识别是人机对话系统中一项重要的研究内容,其任务本质是对用户输入的语句自动标注一个预

先定义好的类别标签^[6]。以往针对用户意图识别的研究主要集中在利用支持向量机(Support vector machine, SVM)、条件随机场(Conditional random field, CRF)等传统机器学习技术,并结合 n-gram 等语言模型的方法^[7-8]对用户意图进行分类,这些方法都是采用较为复杂的人工特征,如语法信息、韵律信息和词汇信息等,较难捕捉文本的深层语义信息。

随着深度学习技术的不断发展,越来越多的研究者开始利用深度学习方法解决自然语言处理领域中的问题,基于深度学习技术的用户意图识别任务也开启了新的发展阶段。Liu 等^[9]使用基于注意力的循环神经网络进行意图识别,用户语句中每个词在 RNN 中都有对应的隐藏状态,利用这些隐藏状态生成最后的意图类别。Ding 等^[10]受 Collobert 等^[11]提出的 CNN 网络结构的启发提出了基于卷积神经网络(CNN)的消费意图挖掘模型(CIMM)来识别用户的消费意图。CIMM 模型首先用一个卷积层将一个上下文窗口中的每个词映射成相应的上下文特征向量,接着用一个最大池化层提取上下文特征向量中的最大值,将池化之后的向量拼接形成句子级的特征向量,最后将句子级特征向量作为前馈神经网络的输入进行一系列的非线性变换,进而实现了消费意图分类。此外,还有一些学者尝试使用深度置信网络(deep belief network, DBN)及深度凸网络(deep convex network, DCN)^[12-13]进行了用户意图识别。

1.2 字符级神经网络

相关研究表明,在文本分类任务中可以通过字符嵌入来得到更加丰富的文本语义表示^[14]。一种方法是利用外部词汇如 WordNet 来丰富单词嵌入,另一种方法是将字符嵌入融入到词嵌入中。Zhang 等^[15]设计了字符级文本分类的卷积神经网络模型,与传统的基于词的卷积神经网络模型及循环神经网络相比取得了更好的效果。Kim 等^[16]于 2015 年也提出使用字符级卷积神经网络提取字词信息,其输出被用作循环神经网络语言模型(RNN-LM)的输入,在词法更丰富的西班牙语数据集上,性能优于传统的词级 LSTM,同时参数也减少了 60%。Santos 等^[17]用英文短文字符序列作为处理单元,分别学习文字的单词和句子特征,以此提高短文分类的准确性。显然,从字符层面进行文本嵌入提取出抽象的文本表示的好处是不需要使用预训练好的词向量等信息,可以尽可能保存原文信息,自然地学习诸如拼写错误和表情符号之类的异常字符组合。此外,字符级神经网络还可以较容易地推广到其

他语言。

2 意图识别模型

2.1 意图识别框架

本文提出的意图识别方法主要分为两个层次:基于敏感词规则的意图识别和基于字符级 RCNN 模型的意图识别。具体的意图识别流程如图 1 所示。

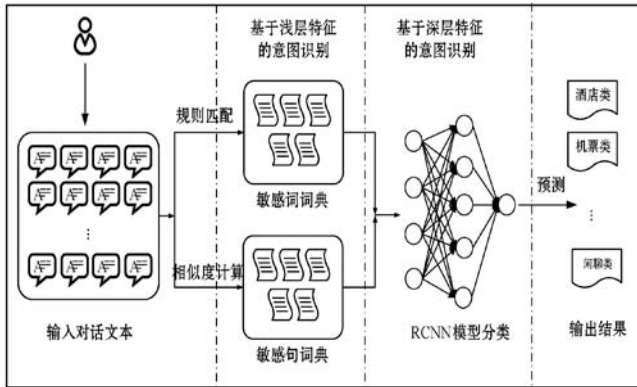


图 1 意图识别流程图

1) 基于敏感词规则的意图识别。利用句子类别词比较明显的数据集统计分析各个类别的敏感词,结合外部网络资源,搜集个别类别的常见关键词,如邮件类、机票类、股票类等,经人工筛选后构建敏感词词典,具体数据样例如表 1 所示;同时还可以针对每个类别构建敏感句词典。意图识别过程首先利用敏感词典,并通过规则及相似度匹配策略对特征明显的对话进行意图识别。

表 1 敏感词表

类别	敏感词
机票类	航班 机场
电话类	打电话 通信录
电影院类	中影 电影院
股票类	走势 涨幅
音乐类	听歌 来首歌 点歌
比赛类	比赛 赛事 赛程

例如用户输入“给张三打电话”、“帮我查询下午去北京的航班”,通过敏感词表中的敏感词“打电话”、“航班”,可以将其归属到电话类及机票类。

2) 基于字符级 RCNN 模型的意图识别。将前一层未能识别的类别特征不明显的对话输入到字符级 RCNN 模型进行分析,得到最终的意图识别结果。

2.2 字符级 RCNN 模型

对用户意图进行分类,首先要对数据进行预处理,

包括字词切分、停用词去除等步骤。传统的分类模型通常选择文本自动分词后的词特征,而基于分词的词语集合构建的文本分类词特征集经常由于分词出错,容易导致模型学习到错误的特征,会大大降低模型的准确率。

例如使用分词工具对“宫保鸡丁怎么做”和“富春山居图电影”这两句话进行分词,分词结果为“宫保鸡/丁/怎么/做”和“富春山/居图/电影”,其中“宫保鸡丁”错误分成“宫保鸡/丁”,“富春山居图”错误分成“富春山/居图”。以上例子采用传统的 1-gram 分词结果进行特征选择时,候选特征会从 {宫保鸡, 丁, 怎么, 做}, {富春山, 居图, 电影} 这些结果选择,显然此时分类器基于这个词语集合会学习出错误的关键词特征,并最终影响模型的实验结果。

因此,本文进一步提出基于字符级的 RCNN 模型。使用神经网络处理数据时需要将其表示为分布式向量,因此模型首先输入用户语句的字符嵌入,通过 LSTM 层后将其输出加入到卷积层,然后卷积层的输出通过最大池化层及全连接层汇集到较小的维度,输出类别标签,模型具体结构如图 2 所示。

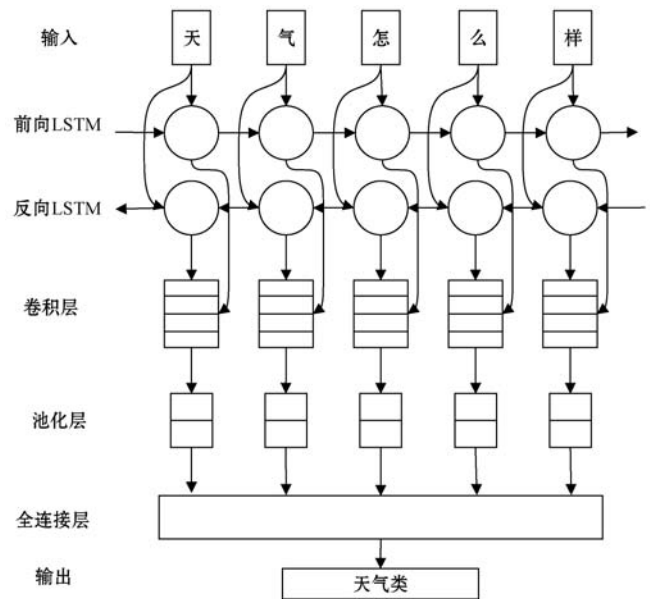


图 2 字符级 RCNN 模型图

1) LSTM 层不仅保存当前的信息,还可以存储任何以前的信息,所以 LSTM 层可以为原始输入生成新的编码。它由一系列重复的时间模块组成,每个模块包含:遗忘门、输入门、输出门和一个记忆细胞结构 $C_t \in \mathbf{R}$ 。输出是 $h_t, c_t, w_f, w_g, w_o, w_i$ 为 LSTM 权重矩阵, b_f, b_g, b_o, b_i 为 LSTM 的偏置, σ 为激活函数 sigmoid, \odot 为点乘操作。在 LSTM 层中的更新方式如下:

$$i_t = \sigma(w_i x_t + u_i h_{t-1} + b_i) \quad (1)$$

$$f_i = \sigma(w_f x_i + u_f h_{i-1} + b_f) \quad (2)$$

$$o_i = \sigma(w_o x_i + u_o h_{i-1} + b_o) \quad (3)$$

$$g_i = \tanh(w_g x_i + u_g h_{i-1} + b_g) \quad (4)$$

$$c_i = f_i \odot c_{i-1} + i_i \odot g_i \quad (5)$$

$$h_{i-1} = o_i \odot \tanh(c_i) \quad (6)$$

2) 卷积层的目的是应用卷积运算,提取任意输入矩阵 \mathbf{W} 隐含高阶特征的过程。通过不同高度卷积核 $C^{l,a}$ 在 \mathbf{W} 上滑动计算,得到不同卷积核的卷积序列,卷积序列构成特征面 $h^{l,a}$,其中 l 表示卷积核高度, a 表示文本向量维度。 $h^{l,a}$ 实际是利用输入矩阵 \mathbf{W} 与 $C^{l,a}$ 进行内卷积的结果集,再加上偏置 $b^{l,a}$ 得到的结果,表示为:

$$h^{l,a} = f(X \times C^{l,a} + b^{l,a}) \quad (7)$$

式中: f 是激活函数 \tanh ,用于对卷积结果作平滑。其目的在于为神经网络引入非线性,确保输入与输出之前的曲线关系。

3) 池化层就是对高维的特征面集合进行降采样操作,进行特征的二次提取。池化操作一般有两种:平均池化和最大池化。本文采用的是最大池化,也就是选取每个特征中的最大值,将池化得到的所有最大值拼接起来,得到一个特征向量。

4) Dropout 是防止神经网络过拟合的技巧,其主要思想是在训练过程中临时从网络结构中以概率 p 随机移除部分神经元以及对应的输入输出权重。

最后使用 softmax 函数进行分类得到意图识别结果。

3 实验及结果分析

为了验证所提方法在用户意图识别中的可靠性,本文分别在两个数据集上进行了多组对比实验,并对实验结果进行了错误分析。

3.1 数据集

针对用户意图识别研究,目前没有标注好大规模数据集,因此本文选取的两组数据集分别为 SMP2017 (中国社交媒体处理大会) 人机对话技术评测中用户意图分类的数据和从百度知道上抓取的问答数据。Broder^[18] 曾提出根据用户行为对用户的意图进行分类,分为导航类、事务类、信息类。在此基础上,SMP2017 所标注的类别将其进行细分为 31 类,基本可以涵盖用户的所有意图。因此,可以使用用户行为类别对用户意图进行表示。数据的样例如表 2 所示,两种数据集规模如表 3 所示。

表 2 数据样例

输入语句	类别
你好啊,很高兴见到你	闲聊
我想订一张去北京的机票	飞机票
今天太原的天气怎么样	天气
李白的静夜思	诗句

表 3 数据集规模

数据集	Smp2017	百度知道
训练	2 290	3 224
验证	769	1 224
测试	667	576
类别数	31	11
文本平均长度	5.02	6.25

3.2 基线模型

为了从不同的角度来验证本文提出的模型在用户意图识别任务中的适用性和有效性。实验采用了几个典型文本分类模型作为基线。

N-gram + SVM: 使用 n-gram 来提取输入特征,并使用 SVM 作为分类器,得到意图识别结果。

Char-CNN: 使用字符级特征作为神经网络模型的输入,通过 CNN 作为分类模型得到意图识别结果。

CNN-LSTM: 使用词级特征作为卷积神经网络输入,提取局部代表性特征,然后按照卷积先后顺序重新组合,再依次输入 LSTM 中进行上下文学习,得到意图识别结果。

Text-GCN^[19]: 在整个语料集上构建图,将词及文档作为图中的节点,利用共现信息重构边,将用户意图识别看作节点分类,以此得到预测结果。

3.3 评价指标

本文采用准确率 (Precision, P)、召回率 (Recall, R) 和 F 值 (F-measure, F) 作为实验评价指标。假定 a 表示被正确分到某类的文档数, b 表示错分到该类的文档数, c 表示属于该类但被错分的文档数。则准确率、召回率和 F 值可分别按照如下公式计算:

$$P = \frac{a}{a+b} \quad (8)$$

$$R = \frac{a}{a+c} \quad (9)$$

$$F = \frac{2 \times P \times R}{P + R} \quad (10)$$

3.4 实验设置及结果

基于 SVM 的意图识别模型选取 1-gram、2-gram 和

3-gram 的组合作为特征集。

对于其他基于深度学习的意图识别模型,采用 skip-gram 模型预先训练嵌入大小为 64 的字嵌入,将 LSTM 的隐藏状态大小设置为 100, CNN 的过滤器宽度设置为 (2,3,4,5),每个过滤器大小为 25。使用 Adam 作为优化器,并将学习率设置为 0.01,丢失率为 0.5,批量大小为 64。

从表 4 可以看出,本文提出的融合敏感词规则和字符级 RCNN 模型的意图识别方法优于所有基线,在两个数据集上的 F 值分别达到了 91.2% 和 90.7%,可以较好地实现用户意图识别任务。在没有使用分词、句法分析等辅助标注工具下,模型能够自动抽取出文本深层次特征,在实际应用中不仅可以简化繁琐的特征抽取过程,而且避免了词法、句法分析的错误传导,减轻了特征选择对识别结果的影响。

表 4 基于两种数据集的实验结果 %

数据集	Smp2017			百度知道		
	P	R	F	P	R	F
Text-GCN	82.5	82.9	82.7	85.3	86.4	85.9
n-gram + SVM	91.7	89.1	89.3	91.0	88.8	89.9
Char-CNN	90.8	88.7	88.9	88.2	87.1	87.7
CNN-LSTM	87.9	81.3	83.5	85.3	86.4	85.9
Our Model	93.6	89.6	91.2	93.1	89.1	90.7

3.5 错误分析

在实验过程中发现,在对 SMP2017 数据进行分词处理时,切分正确率为 80%,且存在较多未登录词切分错误,从而一定程度干扰最后对用户意图的分类效果,因此基于词级的 CNN-LSTM 模型结果低于本文模型,同时是可见的。Text-GCN 表现不好的原因可能在于本任务中短文本较多,导致 GCN 构图时文档-词、词-词之间边比较少,同时 GCN 忽略词的顺序,导致学习的文档向量及词向量区分度不大。

为了进一步分析本文模型的性能,将其预测结果输出进行错误分析,列出了部分具有代表性的分类错误样本,如表 5 所示。

表 5 意图识别错误的部分样本

输入文本	数据集中标注类别	输出类别
今天什么天气怎么样	Chat	weather
我想看电影肖申克的救赎	Video	Cinemas
三安光电	Stock	Video
爸爸去哪几点结束啊	Epg	Chat
2 立方	Calc	Video
可以帮买火车票吗	Chat	Train

经过统计将预测错误样本分为三类:

1) 测试数据样本类别标注错误导致预测错误。表 4 中“今天什么天气”、“可以帮买火车票吗”等属于样本类别标注有误。结合训练数据与类别真实含义来看,模型输出的意图识别结果更准确一些。

2) 样本语义多样性导致预测错误。例如“我想看电影肖申克的救赎”、“爸爸去哪几点结束啊”等样本属于多个类别。

对于“我想看电影肖申克的救赎”,如用户本意是想在线播放,则前一句应属于 video;若是想去影院看,则应属于 cinemas。在这种情况下,仅根据文本的信息无法做出可靠的正确分类。因此对于类似对话语句,在实际应用中需结合额外的场景信息或用户的上下文信息来作出判断。

3) 数据集过少导致预测错误。由于训练样本数据规模较小,而某些类别有大量的领域相关词汇,给定的训练样本无法覆盖该类别的所有语义场景。因而导致模型无法分清“三安光电”是股票还是视频,“2 立方”是数学知识还是视频。

针对这类错误,提升训练样本的数量是一种有效的方法,在此基础上还可以引入类别相关知识,通过构建股票名表、常见数学名词表等知识库可以达到精确的意图识别。

4 结 语

针对人机对话系统中用户对话常为短文本的特点,本文将用户意图识别任务转化为短文本分类任务,并在此基础上提出了融合敏感词规则和字符级 RCNN 模型的用户意图识别方法,分别利用多个基线模型在两类数据集上进行了对比实验,验证了本文提出方法的有效性。接下来我们将在更加开放的领域和更大规模的数据集上进行测试,同时考虑 attention 机制在其中的影响,并扩展基于字符级特征的思想,探索在其他问题上的应用。

参 考 文 献

- [1] Lin R, Costello C, Jankowski C. Enhancing Chinese intent classification by dynamically integrating character features into word embeddings with ensemble techniques[EB]. arXiv: 1805.08914, 2018.
- [2] Aggarwal C C, Zhaic X. A survey of text classification algorithms[M]//Mining Text Data. Berlin: Springer, 2012: 163-222.
- [3] Metzler D, Dumais S, Meek C. Similarity measures for short

- segments of text [M]//Advances in Information Retrieval. Springer Berlin Heidelberg, 2007:16–27.
- [4] Yang L, Qiu M, Qu C, et al. Response ranking with deep matching networks and external knowledge in information-seeking conversation systems [C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. ACM, 2018: 245–254.
- [5] Tur G, Hakkani-Tür D, Heck L. What is left to be understood in ATIS? [C]//Spoken Language Technology Workshop. 2011.
- [6] Qu C, Yang L, Croft W B, et al. Analyzing and characterizing user intent in information-seeking conversations [C]//SIGIR '18, July 8–12, 2018, Ann Arbor, MI, USA, 2018.
- [7] Olney A, Louwerse M, Matthews E, et al. Utterance classification in auto tutor [C]//Proceedings of the HLT-NAACL 03 workshop on Building educational applications using natural language processing—Volume 2, 2003.
- [8] Hutto C J, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text [C]//Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, At Ann Arbor, MI, 2014.
- [9] Liu B, Lane I. Attention-based recurrent neural network models for joint intent detection and slot filling [EB]. arXiv: 1609.01454, 2016.
- [10] Ding X, Liu T, Duan J, et al. Mining user consumption intention from social media using domain adaptive convolutional neural network [C]//Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015:2389–2395.
- [11] Collobert R, Weston J, Bottou L, et al. Natural language processing (almost) from scratch [J]. Journal of Machine Learning Research, 2011, 12(7):2493–2537.
- [12] Zhang B, Kou Y, Meng W U, et al. Close-range air combat situation assessment using deep belief network [J]. Journal of Beijing University of Aeronautics & Astronautics, 2017, 43(7):1450–1459.
- [13] Deng L, Hasegawajohnson M, He X. Random features for kernel deep convex network [C]//IEEE International Conference on Acoustics. IEEE, 2013.
- [14] Tur G, Li D, Hakkani-Tür D, et al. Towards deeper understanding: Deep convex networks for semantic utterance classification [C]//IEEE International Conference on Acoustics. 2012.
- [15] Zhang X, Zhao J, Lecun Y. Character-level convolutional networks for text classification [C]//International Conference on Neural Information Processing Systems. MIT Press, 2015: 649–657.
- [16] Kim Y, Jernite Y, Sontag D, et al. Character-aware neural language models [EB]. arXiv:1508.06615, 2015.
- [17] Santos C N, Gatti M A. Deep convolutional neural networks for sentiment analysis of short texts [C]//Proceedings of International Conference on Computational Linguistics. 2014: 69–78.
- [18] Broder A Z. A taxonomy of web search [J]. ACM SIGIR Forum, 2002, 36(2):3–10.
- [19] Yao L, Mao C, Luo Y. Graph Convolutional Networks for Text Classification [C]//33rd AAAI Conference on Artificial Intelligence, 2018.

~~~~~  
(上接第 155 页)

## 参 考 文 献

- [1] Li J, Wen Z, Li S, et al. Novel tile segmentation scheme for omnidirectional video [C]//2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016: 370–374.
- [2] Misra K, Segall A, Horowitz M, et al. An overview of tiles in HEVC [J]. IEEE journal of selected topics in signal processing, 2013, 7(6): 969–977.
- [3] Skupin R, Sanchez Y, Hellge C, et al. Tile based HEVC video for head mounted displays [C]//2016 IEEE International Symposium on Multimedia (ISM). IEEE, 2016: 399–400.
- [4] Zare A, Aminlou A, Hannuksela M M, et al. HEVC-compliant tile-based streaming of panoramic video for virtual reality applications [C]//Proceedings of the 24th ACM international conference on Multimedia. ACM, 2016: 601–605.
- [5] Corbillon X, Simon G, Devlic A, et al. Viewport-adaptive navigable 360-degree video delivery [C]//2017 IEEE international conference on communications (ICC). IEEE, 2017: 1–7.
- [6] Hosseini M, Swaminathan V. Adaptive 360 VR video streaming based on MPEG-DASH SRD [C]//2016 IEEE International Symposium on Multimedia (ISM). IEEE, 2016: 407–408.
- [7] Fan C L, Lee J, Lo W C, et al. Fixation prediction for 360 video streaming in head-mounted virtual reality [C]//Proceedings of the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video. ACM, 2017: 67–72.
- [8] Sullivan G J, Ohm J R, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard [J]. IEEE Transactions on circuits and systems for video technology, 2012, 22(12): 1649–1668.
- [9] Chen Z, Li Y, Zhang Y. Recent advances in omnidirectional video coding for virtual reality: Projection and evaluation [J]. Signal Processing, 2018, 146: 66–78.
- [10] 董振江, 张东卓, 黄成, 等. 虚拟现实视频处理与传输技术 [J]. 电信科学, 2017, 33(8): 45–52.
- [11] 杨帆. 面向头盔显示的全景视频播放系统 [D]. 上海: 复旦大学, 2016.