Tuijin Jishu/Journal of Propulsion Technology
ISSN:1001-4055
Vol. 44 No.3 (2023)

A Review of Machine Translation Techniques and Proposal to Automatic Translation of Videos

Shenbagaraj Ramachandran1* and Sailesh Iyer1

1 Department of Computer Science and Engineering, Rai University, Ahamedabad, India

Correspondence to: Shenbagaraj R, Research Scholar, Department of Computer Science and Engineering

Rai University, SH144, Village – Saroda, Dist – Ahamedabad – 382260 - Gujarat, India

shenbagaraj.ramachandran@raiuniversity.edu

***Abstract*—** Translation of videos including propulsion technology videos has been one of the toughest jobs to do during the eLearning development process. We have tried it earlier by involving mostly manual work and found it to be tiresome. After the fast evolution of Machine learning and Natural language processing techniques, it is feasible now to do the translation involving and inculcating intelligent techniques in the process. We have reviewed the technologies and techniques and presenting it in this paper.

Keywords— *machine translation, adaptive eLearning, Machine learning, animated videos*

Introduction

The importance of the language option in Adaptive learning is already emphasized by us in the paper [1]. This option of learning using the native language is being requested by several colleges, universities, and international clients as they see the benefits of usage of native language. But some universities may prefer teaching using an international standard language like English. So, while providing eLearning solutions for these universities, having materials in both English and Native languages will be beneficial. But translating the eLearning materials especially eLearning videos from English to their native language is a challenging task. Automating Translation is one of the hardest problems as many contextual and expressive nativity of different languages [2].

Earlier it involved the skills, time and exhaustive teamwork of human translators, voice artists, and subject matter experts to convert a video from English to Native languages. This traditional process is expensive and time consuming. The effort needed is going to be intensified if we need translation in multiple languages. Recently, Machine learning based translation is a feasible solution that is being explored as the accuracy has improved [2].

In this paper, animated English videos are machine translated in to different languages used across the world by applying machine learning cloud services. This project was proposed by us based on the need and demand to convert and use the existing English videos.

Section 2 discusses the related research work. Section 3 describes the proposed methodology while section 4 presents the expected results. Section 5 concludes the paper suggesting an implementation considering the intelligent automation and production time.

Related Work

Researchers have been exploring Machine Translation for many years and have found that Machine Translation has been more accurate after the recent evolution of the machine learning and deep learning techniques. Here, we are listing related works which were the base for our proposed methodology and inspiration for taking this research work.

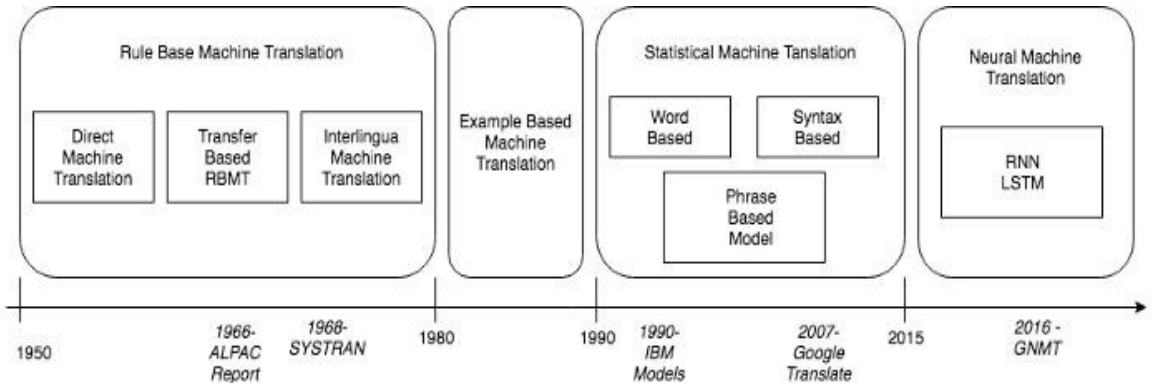
In 1799, a Danish scientist Christian Kratzentein was the first to work on translation systems. Thereafter with lot of research, many improvements were gradually made. Inspired by IBM founder Thomas Watson Sr, IBM came up with a translation system in the year 1927 itself for which the ground work started from 1920s [3]. In 1949, during the Second World War based on the information theory and successes in code breaking, researcher Warren Weaver at Rockefeller Foundation proposed for machine based translations. Then many universities in US started research on Machine Translation [4]. In 1954, Georgetown-IBM experiment started which gained much media and public interested as the IBM 701 computer automatically translated 60 Russian sentences into English for the first time in history. Unfortunately the ALPAC report stated that machines cannot compete with the human translation quality and suggested to stop funding for Machine Translation research. But several researchers continued their study on automatic machine translations and kept the ball rolling. Most of these researchers concentrated on limited language pairs with limited inputs and used rule based engines. By the 1980s, most of machine translation engines were based on Main Frame technologies In 1981, P.Rubin and Thomas Baer came up with an idea of an articulatory synthesizer for perpetual research [4].

The initial effort in India regarding machine translation was tried by RMK Sinha et al at IIT Kanpur [5]. This effort involved human involvement as a post-production work. This was reduced to some extend by developing an another application called “Anusaaraka”[6]. Only Indian languages were considered in these applications.

The Evolution of Machine Translation can be divided in to

1. Rule Base Machine Translation
2. Example Based Machine Translation
3. Statistical Machine Translation
4. Neural Machine Translation

Figure 1: Evolution of Machine Translation



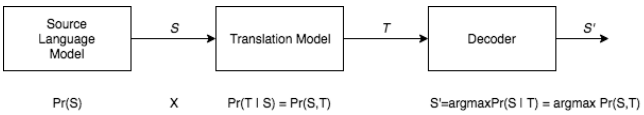
Rule based Machine Translation and Example based Machine Translation has been already discussed so we will list the literature review on Statistical Machine Translation and Neural Machine Translation.

Statistical Machine Translation

In 1990, Brown et al proposed the use of statistical methods in Machine Translations. They proposed a translation process where the source text is partitioned into a set of fixed locations, then the glossary is used to select the set of fixed locations to create a sequence, and finally words in target fixed locations are rearranged to

form a target sentence. They successfully developed the statistical techniques for automatic glossary creation and arrangement of target word sequences but failed to provide examples for translated sentences [7].

Again in 1993, Brown et al described a series of five statistical models for the translation process as shown in the figure below



.

Figure 2: Process and Translation Model of Brown et al

And gave algorithms for estimating the parameters of these models given a set of bilingual pairs of sentences. These models were later considered as the IBM alignment models. They defined the concept of word-by-word alignment between the pair of bilingual sentences. Their algorithm assigned a probability to each of these word-by-word alignments for any given pair of sentences. Though their research was confined to smaller English and French translations but it was a considerable improvement to the alignment of word-by-word relationships in the pair of sentences [8]

In 1996, Vogel et al described a new model for word alignment in the Statistical Machine Translation using first-order Hidden Markov Model as it solved the time alignment problem for speech recognition. The main idea behind the model was to make the word-by-word alignment probabilities depend on the alignment positions rather than the absolute positions. The HMM-based model produced translation probabilities on par with the mixture alignment model and position alignments were much smoother in HMM-based model [9].

Och et al described a method to determine bilingual word classes to be used in Statistical Machine Translation. They developed an optimization criterion based on the maximum likelihood approach and further described a clustering algorithm as shown in the figure below.

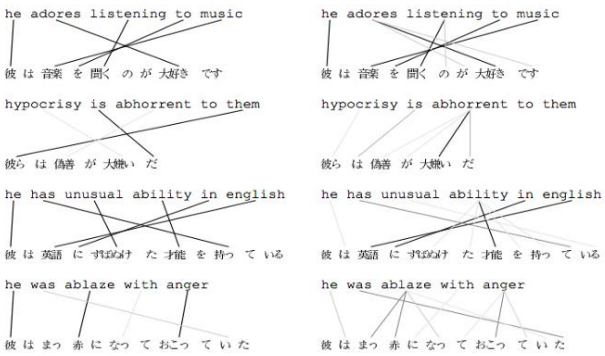


Figure 3: Illustration of clustering algorithm by Och et al

The results of their experiments showed that the usage of bilingual word classes improved the statistical machine translations significantly [10].

A syntax-based statistical translation model was proposed by Yamada et. al (2001). Their model transformed a source-language parse tree into a target-language string by applying stochastic operations at each node. Those operations captured the linguistic differences such as word order and case marking. The model produced word alignments which were better those produced by IBM Model 5 [11].

A novel phrase-based translation model and decoding algorithm was proposed by Koehn et al (2003) which enabled them to evaluate and compare several previously proposed phrase-based translation models. They

designed a uniform framework to compare different other translation models. The model proposed by Koehn et al (2003) was based on the noisy channel model Brown et al (1993) and they used the Bayes rule to reformulate the translation probability for translating a foreign sentence in French into English [12].

Chiang et al (2005) presented a phrase-based machine translation model that used hierarchical phrases – phrases that contained sub phrases. They proposed the use of hierarchical phrases which consisted of both words and sub-phrases to address this problem. Their model was based on a weighted synchronous Context Free Grammar. The model built partial translations using the hierarchical phrases and then combined them serially in a standard phrase-based model. Instead of using the traditional noisy-channel approach, they used a more general log-linear model [13].

Neural Machine Translation

In 2013, a new end-to-end encoder-decoder structure for machine translation was proposed by Kalchbrenner & Blunsom(2013). They introduced a class of probabilistic continuous translation models called Recurrent Continuous Translation Models which were purely based on continuous representations for words, phrases and sentences and did not rely on alignments or phrasal translation units [14].

Sutskever et al (2014) proposed the use of Deep Neural Networks in Sequence to Sequence learning for Machine Translations. Their method used a multi-layered Long Short-Term Memory (LSTM) as shown in the figure 4 to map the input sequence to a vector of fixed dimensions, and then used another deep LSTM to decode the target sequence from the vector.

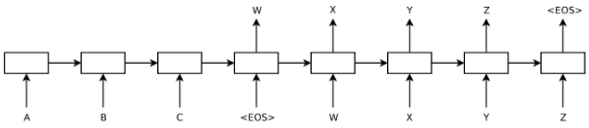


Figure 4: Multi Layered LSTM

Their results showed that Neural Machine Translation system having large deep LSTM with a limited vocabulary can outperform a standard SMT-based system [15].

Bahdanau et al (2014) proposed a method which allowed a model to automatically soft-search for parts of a source sentence that are relevant in predicting a target word, without having to form these parts as a hard segment explicitly. With this approach, they achieved e a translation performance comparable to the existing state-of-the-art phrase-based system on the task of English-to-French translation [16].

Luong et al (2015) proposed two effective classes of attention mechanism, a global approach which always attends to all source words and a local one that only looks at a subset of source words at a time. Their ensemble model using different attention architectures established a new state-of-the-art result in the WMT'15 English to German translation task with 25.9 BLEU points [17].

Jozefowicz et al (2016) experimented with different neural network models on different sizes of corpora, their experiments showed that RNNs can be trained on large amounts of data, and they outperform competing models including carefully tuned N-grams. Their experiments showed that a large, regularized LSTM LM, with projection layers and trained with an approximation to the true Softmax with importance sampling performed much better than N-grams [18].

In the Findings of First Conference on Machine Translation (WMT'16), the neural machine translation systems that participated in the WMT evaluation outperformed phrase-based statistical machine translation system by up to 3 BLEU score (Bojar et. al 2016) [19] .

In the paper [20] Bibekananda Kundu and Sanjay Kumar Choudhury proved that a machine learning based approach will help in machine translation systems while using for English to Bangla conversion.

Mary Priya et al. [21] came up with a statistical framework for translating English documents to Malayalam. However the framework only considers only text documents and consumes a huge amount of memory. The space complexity was overcome by Remya Rajan et al in their published work [22]. While this translation system proposed involves Roman and Unicode files, a morphological analyzer for the Malayalam language was proposed by Rinju O.R et al [23]. It was inferred from their results that a rule-based approach was better than the probabilistic method. Apart from the machine translation efforts done for English to Malayalam there are systems for each and every language like the one proposed by Keerthi Lingam et al. [24] for English to Telugu. This method only focuses on prepositions because of the multiple meanings for the same preposition in Telugu language. This method was also developed based on a rule-based approach. Text in diagrams and flow charts were identified and converted in this method.

Srikar Kashya Pulipaka etal [25] used an open source applications for machine translation, text-to-speech and speech-to-text to convert English language video to some of Indian language audio for visually impaired people and reading handicapped people. Their output is only audio and here we propose to a matching video in multiple languages. Moreover, video with multiple voices are not considered in their work. The case of conversational videos is also not considered in their work.

Most of the related work has been focused and done for text or speech based input and output. Not much research work has been done for the video based translation work for multiple languages across the world. So, we proposed to work on this area as we had a requirement internally to convert the English videos to multiple languages with less human involvement.

Proposed Methodology

There are several Machine Translation applications developed for various language but they are either developed for individual languages or individual multimedia elements or purposes alone. This proposed system is a collective tool where the users can choose the desired language and the intelligent cloud services based on neural translation to convert the English videos to the target language videos. The following steps are proposed to be followed to translate the videos

Getting the transcript from the source video

1. Translate the transcript in to the target language
2. Create corresponding sub-title files
3. Combine all those pieces to get the target video

The following diagram illustrates the process in detail:

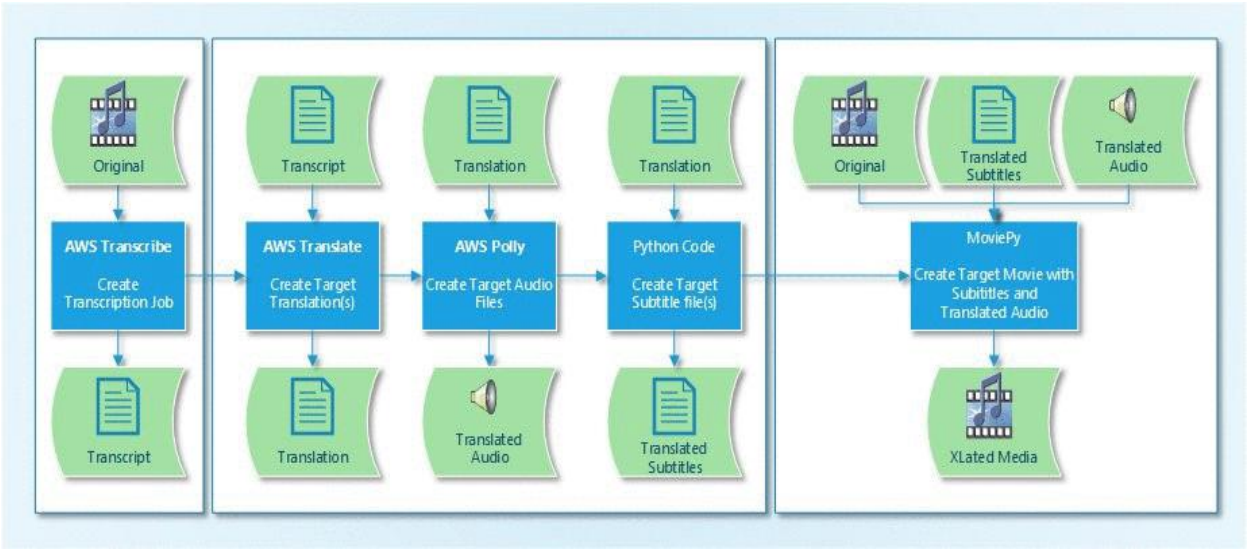


Figure 5: Suggested Automatic Translation Process

Our proposed process begins by creating a transcription from the original video by using any of the intelligent cloud services and transcript is translated to get the translation. For better translation and syncing later, we need to do an intelligent splitting based on the context of the contents in the story board used to create the animated videos. Then the translated script is used by the appropriate intelligent cloud service to generate the translated audio. As the over dubbed videos in which the lips of the characters don't line up with the audio, we treated subtitles and their audio for translations differently compared to the original transcript. We used MoviePy which leverages ImageMagick and FFmpeg to build key functionalities like text titles, animation, audio, and videos. Then, we will replace the audio from the video clip. Next, we will create the subtitles. After the subtitles are created, we will create an array of subtitled clips. Next, we will concatenate all of the clips in to one final clip. Finally. We will write the subtitled video and audio to a new video file.

Expected Result

Video to Video translation system for English to various languages was implemented by using neural based machine translation method. For evaluating the system, 50 videos from various languages were selected. Each of these videos was tested using the proposed methodology and the tool developed. Performance evaluation of the automatic machine translation system is given in the Table 1

TABLE 1. Evaluation of Machine Translation System

nber of Videos to be taken for chine Translation including pulsion technology videos	
uracy	
perts	an Translation Experts

Conclusion

We conclude by saying that machine translation using machine learning is expected to drastically reduce the manual work and increase the productivity of the translation work. Based on this review paper, we decide that Neural based translation techniques will result in accuracy closer to human expert-based translation.

Acknowledgment

We acknowledge our management for allowing us to carry out this proposed research work and motivating us to take up this project work.

References

[1] Shenbagaraj R., Iyer S. (2021) A Smart Education Solution for Adaptive Learning with Innovative Applications and Interactivities. In: Hassanien A.E., Bhattacharyya S., Chakrabati S., Bhattacharya A., Dutta S. (eds) Emerging Technologies in Data Mining and Information Security. Advances in Intelligent Systems and Computing, vol 1300. Springer, Singapore. https://doi.org/10.1007/978-981-33-4367-2_65

[2] AWS Machine Learning Blog, aws.amazon.com/blogs/machine-learning/create-video-subtitles-with-translation-using-machine-learning/, last accessed 2021/07/10.

[3] IBMWebsite, https://www.ibm.com/ibm/history/ibm100/us/en/icons/translation/

[4] Towards data science website,

https://towardsdatascience.com/evolution-of-machine-translation-5524f1c88b25

[5] R.M.K. Sinha, K. Sivaramam, Aditi Agrawal, Renu Jain, Rakesh Srivastava, Ajai Jain, "ANGLABHARTI: A Multilingual Machine Aided Translation Project on Translation from English to

Indian Language", Proc. IEEE, 1995. Available: IEEE https://ieexplore.iee.org/stamp/stamp.jsp?tp=&arnumber=538002&tag=1.

[6] Akshar Bharati, Vineet Chaitanya, Amba P.Kulkarni and Rajeev Sangal, " ANUSAARAKA: Machine Translation in Stages", A Quarterly in Artificial Intelligence, Vol 10, No. 3, July 1997.

[7] Brown et al, "A Statistical Approach to Machine Translation", Computational Linguistics Volume 16, November 2, June 1990

[8] Peter F.Brown, Stephen A.Della Pietra, Vincent J.Della Pietra, Stephen A.Della Pietra, Robert L.Mercer, IBM T.J Watson Research Center, "The Mathematics of Statistical Machine Translation: Parameter Estimation", 1993, Association of Computational Linguistics.

[9] Stephan Vogel, Hermann Ney and Christoph Tillmann, "HMM-Based Word Alignment in Statistical Translation", COLING 1996 Volume 2: The 16th International Conference on Computational Linguistics.

[10] Franz Josef Och, Christoph Till mann, and Hermann Ney, " Improved Alignment Models for Statistical Machine Translation, 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora.

[11] Kenji Yamada and Kevin Knight, "A Syntax-based Statistical Translation Model", Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics.

[12] Philip Koehn, Franz J.Och, Daniel Marcu, " Statistical Phrase-Based Translation" , Proceedings of the 2003 Human Langauge Technology Conference of the North American Chapter of the Association for Computational Linguistics.

[13] David Chiang, " A Hierarchical Phrase-Based Model for Statistical Machine Translation", University of Maryland, College Park, MD

[14] Nal Kalchbrenner and Phil Blunson, "Recurrent Contuinous Translation Models", Proceedings of the 2013 Conference on Empirical Methods on Natural Language Processing.

[15] IIya Sutskever, Oriol Vinyals and Quoc V.Le, "Sequence to Sequence Learning with Neural Networks, Google, Computation and Language, Computer Science, Cornell University

[16] Dzmitry Bahdanau and KyungHyun Cho and Yoshua Bengio, "Neural Machine Translation by Jointly learning to Align and Translate", Published as a conference paper at ICLR 2015.

[17] Minh-Thang Luong, Hieu Pham and Christoper D.Manning, " Effective Approaches to Attention-based Neural Machine Translation", CSE Dept, Stanford University.

[18] Rafal Jozefowicz et al, " Exploring the Limits of Language Modeling", Google Brain.

[19] Ondrej Biojar et al, "Results of the WMT16 Metrics Shared Task", Proceedings of the First Conference on Machine Translation:Volume 2, Shared Task Papers.

[20] Mary Priya Sebastian, Sheena Kurian and G. Santhosh Kumar , " A Framework of Statistical Machine Translator from English to Malayalam,", Proc. 1st Amrita ACMW Celebration on Women in Computing, September, 2010.

[21] Remya Rajan, Remya Sivan, Remya Srinivasan, K.P.Somam, "Rule Based Machine Translation from English to Malayalam", International Conference on Advances in Computing, Control and Telecommunication Technologies, 2009.

[22] Rinju O.R, Rajeev R.R, Reghu Raj P.C, Elizabeth Sherly, " Morphological Analyzer for Malayalam: Probabilistic Method Vs Rule Based Method", International Journal of Computational Linguistics and Natural Language Processing, ISSN 2279-0756, vol. 2, Issue 10, October 2013.

[23] Keerthi Lingam, E. Rama Lakshmi and L.Ravi Teja, "Rule-based Machine Translation from English to Telugu with Emphasis on Prepositions", International Conference on Networks & Soft Computing, 2014, Available [Online] IEEE: https://ieeexplore-ieee-org.libaccess.sjlibrary.org/stamp/stamp.jsp?tp=&anumber=6906669

[24] D S Sharma, R Sangal and J D Pawar, Proc of the 11th Intl. Conference on Natural Language Processing, pages 167-176, Goa, India, December 2014.

[25] Srikar Kashyap Pulipaka et al (2019) , Machine Translation of English Videos to Indian Regional Languages using Open Innovation, IEEE International Symposium on Technology in Society (ISTAS) Proceedings Miriam Cunnigham and Paul Cunnigham (Eds) ISBN:978-1-7281-5480-0