Improving text detection using Augmented Reality for Android-based Smart phones

Arun Prajapati

ME (CE), LJIET Ahmedabad, Gujarat, INDIA

Swarndeep Saket

Assistant Professor PG Department (CE)
LJIET Ahmedabad, Gujarat,INDIA

Abstract

Augmented Reality (AR) is an enhanced version of reality where live direct or indirect views of physical real-world environments are augmented with superimposed computer-generated images over a user's view of the real-world, thus enhancing one's current perception of reality. Text Detection is very useful in day to day life as it make easy to write and understand text by just using the Smart Phones. In this Dissertation work, we are going to improve the User Experience with Optical Character Recognition (OCR) for better performance, as OCR is responsible for text detection. Here we are going to user live Detection and translation of text to improve User Experience Using Augmented Reality for Android-Based Smart-Phones.

Keyword: Augmented Reality; Optical Character Recognition (OCR); Text Detection.

1. Introduction

Augmented reality (AR) is an interactive experience of a real-world environment where the objects that reside in the real-world are "augmented" by computer-generated perceptual information, sometimes across multiple sensory Ahmedabad, Gujaratmodalities, including visual, auditory, haptic, somatosensory, and olfactory.

The overlaid sensory information can be constructive (i.e. additive to the natural environment) or destructive (i.e. masking of the natural environment) and is seamlessly interwoven with the physical world such that it is perceived as an immersive aspect of the real environment.

Optical character recognition or optical character reader, often abbreviated as OCR, is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast).

2. Pervious Work

In recent years, new discriminative features have been proposed for text detection, including the Stroke Width Transform (SWT) and Maximal Stable Extremal Re- gions (MSER), both of which have been used widely. Most characters have a narrow and uniform stroke width, along with clear edges and homogeneous colors. SWT and MSER are designed to capture these properties.

A variety of machine learning techniques have been used for text detection, including unsupervised feature learning, Convolutional Neural Networks, deformable part- based models , belief propagation , and Conditional Random Fields . Bai et al. identify text regions using gradient local-correlation to find edge pairs and estimate stroke width. The relationship between different CCs, colors and shapes are fed into SVM classifiers to detect text.

3. Methodology

AR technology is widely used in all possible areas in the world, there is still some issue to detect and translate the text to user preferable language. So here are making a model which will improve the user experience and process of text detection and translation.

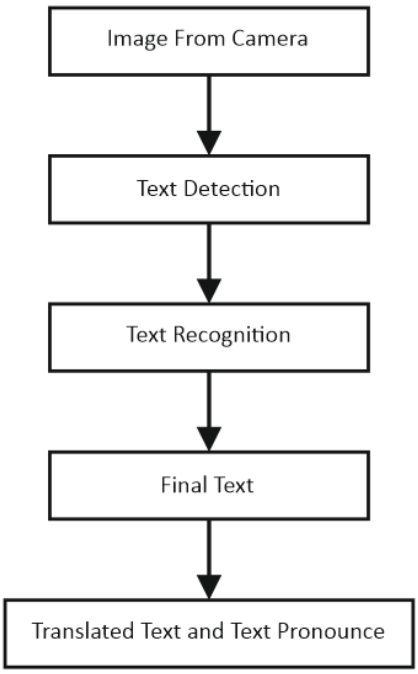

Figure Proposed System

The Proposed System use the camera to take the live image then it will try to detect all the possible text in live image with help of OCR, then text reorganization takes place and finally the text is translated to other Language and displayed near the detected text.

Step 1: Get Image from Camera, The first step is to provide the input to our model, the input is a live image which is provided directly by camera, the live feed is used as input it can be considered as image or video as with are not dealing with still images.

Step 2: Text Detection, using Conventional Text Detection Algorithm Here the Text Detection takes place using OCR and the detected Text is sent for text recognition.

Step 3: Text Recognition, Here Text Recognition algorithm is Applied and possible text are detected.

OCR is generally an "offline" process, which analyses a static document. Handwriting movement analysis can be used as input to handwriting recognition. Instead of merely using the shapes of glyphs and words, this technique is able to capture motions, such as the order in which segments are drawn, the direction, and the pattern of putting the pen down and lifting it. This additional information can make the end-to-end process more accurate. This technology is also known as "on-line character recognition", "dynamic character recognition", "real-time character recognition", and "intelligent character recognition".

Step 4: Final Text, Finally will get all Texts as words in an array.

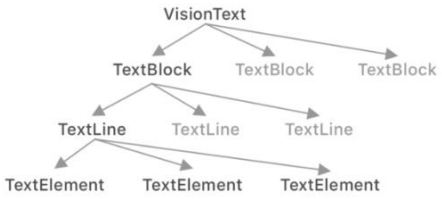

Figure 3.2 Vision Text Structure [23]

Step 6: Translated Text, We will try to translate all the words in array to particular language.

Step 7: Pronouncing Text, Finally the Translated Texts will be pronounced.

Text to Speech synthesis is the process of text transformation into human speech audio using computer. Many Operating System had built the feature for text to speech since 1990s. The advancement in Machine Learning and Artificial Intelligence in recent years results in many new advance voice synthesis technologies such as WaveNet Deep Learning Neural Network from DeepMind.

We will use TTS in Flutter

3.4 Expected Outcome

The proposed system will detect and translate the text on live camera fast and with improved performance.

3.5 Tools and Techniques

Android Studio

Flutter (Framework)

Dart 4. Experiment

We tested our System using Android Smart phone over 300 objects which are printed, like match box, visiting card and other printed material

4.1 Result

we got success above 80% in all experiment we carried out

here are few output captured while experimenting

as you can see we detected Camlin whiteboard ink box's text even from a good distance as the distance totally depend on the quality of camera as now smartphones come with good camera specifications.

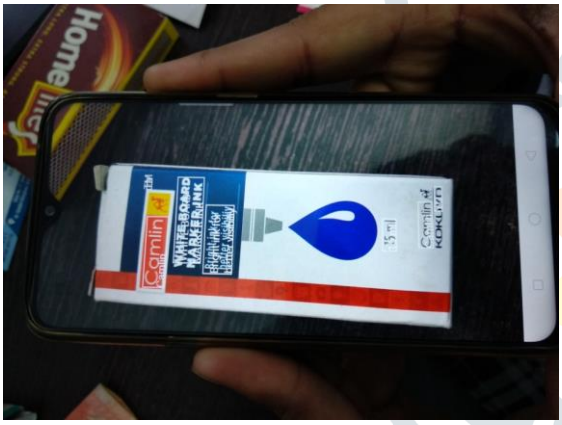

Figure 5.1 Live Text Detection and Display on Camlin White Board Ink Box

Figure 5.2 Live Text Detection and Display on Visiting card

Figure 5.3 Live Text Detection and Display on mobile camera lens Box

5. Conclusion

We have proposed a relatively simple cascaded text detection system that is accurate at the pixel, character and word levels By this method the text detection is improved as well the text processing is done on the same screen which gives the best user experience. It is important to note that we obtain high accuracy using only visual features, without the use of a language model or dictionary. However, detection can probably be improved by integrating character recognition and models for a specified language.

References

[1] Hyung Il Koo, BeomSuKim, Young Ki Baik and NamIkCho (2016) "Fast and Simple Text Replacement Algorithm for Text-based Augmented Reality",IEEE, pp 27-30.

[2] Fedor Borisyuk, Albert Gordo and Viswanath Sivakumar (2018) "Rosetta: Large Scale System for Text Detection and Recognition in Images*", Applied Data Science Track Paper, pp 71-79.

[3] Mu Jung Cho,Anupriya Gagneja, Xiao Yang, Miriam Brinberg, Katie Roehrick, Sagnik Ray Choudhury, Agnese Chiatti, Nilam Ram, Byron Reeves and C. Lee Giles(2018) " Text Extraction and Retrieval from Smartphone Screenshots: Building a Repository for Life in Media " , Association for Computing Machinery, pp 948-955.

[4] Siyu Zhu and Richard Zanibbi (2016) "A Text Detection System for Natural Scenes with Convolutional Feature Learning and Cascaded Classification",IEEE Conference on Computer Vision and Pattern Recognition, pp 625-632.
[5] Satoshi Yonemoto (2014) "A method for Text Detection and Rectification in Real-world Images", 18th International Conference on Information Visualization, pp 374-377.Kajaree Das1, Rabi Narayan Behera2 " A Survey on Machine Learning: Concept, Algorithms and Applications" IJIRCCE Vol. 5, Issue 2, February 2017 P.P 1301 – 1309
[6] MICHAEL CUTTER and ROBERTO MANDUCHI (2017), "Improving the Accessibility of Mobile OCR Apps Via Interactive Modalities", ACM Transactions on Accessible Computing, Vol. 10, No. 4, Article 11. Publication date: August 2017.
[7] Victor Fragoso,Steffen Gauglitz, Shane Zamora, Jim Kleban, Matthew Turk (2010),"TranslatAR: A Mobile Augmented Reality Translator",978-1-4244-9497-2/10/$26.00 ©2010 IEEE, pp 497-502
[8] Jorge Martín-Gutiérrez, Peña Fabiani, Wanda Benesova,María Dolores Meneses,Carlos E. Mora (2015)," Augmented reality to promote collaborative and autonomous learning in higher education"Computers in Human Behavior 51, pp 752–761
[9] Marc Petter, Victor Fragoso, Matthew Turk, Charles Baur (2011),"Automatic text detection for mobile augmented reality translation", 978-1-4673-0063-6/11/$26.00 c 2011 IEEE
[10] Rohit Saluja, Devaraj Adiga, Parag Chaudhuri, Ganesh Ramakrishnan and Mark Carmanq (2017), " Error Detection and Corrections in Indic OCR using LSTMs", 2017 14th IAPR International Conference on Document Analysis and Recognition, pp 17-22 2017 IEEE,
[11] https://dataaspirant.wordpress.com/2014/09/19/supervised-and-unsupervised-learning/ Last Accessed on 3/10/2018 9:50 pm.
[12] https://www.3pillarglobal.com/insights/augmented-reality-introduction-and-its-real-world-uses Last Accessed on 3/10/2018 9:56 pm.
[13] http://reinforcementlearning.ai-depot.com/ Last Accessed on 3/10/2018 10:06 pm.
[14] https://www.researchgate.net/figure/Figure-1-Block-Diagram-of-the-Proposed-OCR-System-Model_fig1_260405352 Last Accessed on 3/10/2018 10:16 pm.
[15] https://en.wikipedia.org/wiki/Optical_character_recognition Last Accessed on 15/2/2019 9:36 pm.
[16] https://pdfs.semanticscholar.org/8fa4/5b60e78b3d7cc26271dfa79baea66e7d13ce.pdf Last Accessed on 09/10/2018 11:40 pm.
[17] https://opencv.org/about.html ast Accessed on 15/10/2018 1:03 am.
[18] https://opensource.google.com/projects/tesseract Last Accessed on 28/10/2018 10:06 pm.
[19] https://en.wikipedia.org/wiki/Tesseract_(software) Last Accessed on 28/10/2018 10:06 pm.
[20] https://en.wikipedia.org/wiki/Android_Studio Last Accessed on 13/11/2018 8:36 pm.
[21] https://en.wikipedia.org/wiki/Flutter_(software) Last Accessed on 15/2/2019 9:34 pm.
[22] https://en.wikipedia.org/wiki/Dart_(programming_language) Last Accessed on 15/2/2019 9:36 pm.
[23] https://medium.com/flutter-community/flutter-mlkit-8039ec66b6a Last Accessed on 20/2/2019 9:36 am.