

Koraku Word Recognition Using Correlation and Find it's Appropriate (Antonyms) In Marathi

Arvind A. Tayade

G. S. Science, Arts and
Commerce College, Khamgaon(M.S.)
arvindtayade40@gmail.com

Anand M. Khandare

anandkhandare99@gmail.com

Abstract: *The vast field of computer science face the number of challenges and overcome the problems of society via technology. Computer vision deals with number of problems and solve those problems efficiently. Inter collaboration is a thrill of research, the part of those things we have to work in new domain world. In this paper, we worked on the Koraku word recognition and find its Marathi correspondent (antonyms). To perform those things, we use correlation method for the recognition of koraku words.*

I. INTRODUCTION

Now a days image processing is applicable in every field. Small modules belongs to some application of image processing like mobile biometric recognition, Face Biometric, Fingerprint Biometry. Such types of application currently popular in our society. One of them, we introduce in our propose technique to recognize the koraku language text and find it's appropriate meaning in Marathi word dictionary.

Koraku language is an ancient language which is basically use in Melghat (Dist -Amravati) area. Koraku language is a communication medium of Koraku jamat, which is derived from the combination of partially Hindi and Marathi, such as 'KAMAY' in koraku language means 'KAM' in Marathi.

According to current research of digital image processing, research work is done in three parallel ways such as low level, middle level and high level image processing. Text recognition process performed in both middle level as well as high level image processing. But as per our point of view, those techniques are more critical and time consuming and complexity of algorithm. According to [1] a method may include receiving a text input in a script and segmenting the text input into one or more graphemes. Each of the one or more graphemes may be split into one or more recognition units based on one or more recognition unit identification criteria associated with the script. Next, a text recognition system may be trained using the recognition units. The basic methodology of image word recognition [4] is start with raw image segment of text. After segmentation of text, segment the word of text. When two level segmentation is completed then recognize the character and finally contextual verification of that text. To complete this

long term process system require more time for recognition of word. Some researchers proposed [2] the some new approach which is started from Image acquisition, noise removal, Binary conversion, morphological operations, connected components, Features extraction, segmentation, matching characters, performance evolution etc. But according to application point of view this process is time consuming.

The digital image correlation (DIC) technique provides displacements and strain maps on deformed surfaces. The correlation is possible only if the surface has a random texture, such as a black and white speckle [5].

II. OBJECTIVE OF RESEARCH WORK

1. Proposed new technique for word recognition.
2. Recognition of Koraku language word using correlation method.
3. Create dictionary of koraku and Marathi text.
4. Create API for recognition and find conversion of Koraku language words.

III. PROPOSED TECHNIQUE

- A. **Image Acquisition :** In proposed method, Koraku language word image is our input image, we are use google input** for acquire the image from the client side and save this image in monochrome bitmap image. This work is the basic work for manipulating with the koraku language text. So simple experimentation perform on the image. For image acquisition, the size must be 512 x 256 pixel with monochrome bit map image.



Figure-1 Sample Koraku Word Image

- B. **Correlation of Input Image and Database Image :** After acquisition of input image, compare the correlation of input image and image present in our database and acquire the maximum value of correlation

i.e. the value of correlation in between 0 to 1. If the value of correlation get more than 0.5 or greater found in whole database, this is the best combination in database. If correlation is 1 this means that image recognize is 100%. It may be possible to down in some similar type of words.

Following equation calculates the correlation of two images. Range of value starts from 0 to 1. In this equation, A and B are a two dimension matrices. One important thing in correlation, the dimension of two matrix must be same.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right) \left(\sum_m \sum_n (B_{mn} - \bar{B})^2\right)}}$$

Where, \bar{A} = mean2 (A), and \bar{B} = mean2 (B)

In our work, A and B are two dimensional images. This operation cannot perform on the color image. But in our era of research work, in word recognition there may be no need

Table -1 Correlation of Koraku word

	सेने	हजे	मांडीये	आयुमे	त्याराकू	..
सेने	1	0.5001	0.2662	0.3778	0.325	..
हजे	0.5001	1	0.2137	0.2595	0.211	..
मांडीये	0.2662	0.2137	1	0.2956	0.322	..
आयुमे	0.3778	0.2595	0.2956	1	0.239	..
त्याराकू	0.3246	0.2107	0.3222	0.2385	1	..
.	1..

- C. **Koraku database** : We have create our own standard database with 20 Koraku Words. The pixel size of database is 512 X 256 created in paint*. Arial Unicode MS Font style used for database creation. Save this image in monochrome bitmap image. Monochrome bitmap image show the single color in spatial space.
- D. **Marathi database** : In Marathi words database contents 20 corresponding words of Marathi language the size of Marathi word image is 1024 X 256 pixel. This image set also created in paint* using google input tool**.
- E. **Output image**: After compare the input image and database image correlation, find the high correlated image in database and display its corresponding image from Marathi dataset.

IV. CONCLUSION

Current text recognition techniques are complex to execution, they will carry number of operation for text recognition. Apart from those techniques, our proposed technique is very good in the case of word verification. This work is also use full to Koraku Community students

of color. We cannot treat color as a feature, so it is not a drawback of this technique.

Following table show the correlation of Koraku word image with respect to itself and other koraku word image.

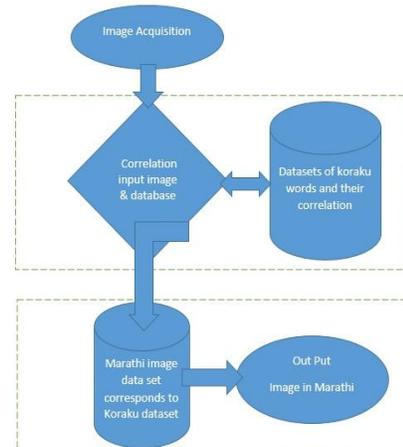


Figure - 2 Proposed Method

for better understanding of Marathi language because of Marathi community relates to Koraku community for many purposes. This work is initial work in this field, number of things are remaining to improve the results.

REFERENCES

- [1] Deselaers, Thomas, et al. "Text recognition based on recognition units." U.S. Patent No. 9,536,180. 3 Jan. 2017.
- [2] Janani, G., V. Vishalini, and P. Mohan Kumar. "Recognition and analysis of Tamil inscriptions and mapping using image processing techniques." *Science Technology Engineering and Management (ICONSTEM), Second International Conference on.* IEEE, 2016.
- [3] Bhat, Nalin, Avinash Kumar Yadav, and Krushna Rajbinde. "Recognition, Formatting In Image Files By Using Image Processing." *Imperial Journal of Interdisciplinary Research* 2.3 (2016): 51-54.
- [4] Elliman, David G., and Ian T. Lancaster. "A review of segmentation and contextual analysis techniques for text recognition." *Pattern Recognition* 23.3-4 (1990): 337-346.
- [5] Flament, Camille, et al. "Local strain and damage measurements on a composite with digital image correlation and acoustic emission." *Journal of Composite Materials* 50.14 (2016): 1989-1996. *<https://support.microsoft.com/en-us/instantanswers/8a779cf7-3362-4948-a65e-e3a78c86f9a8/download-paint>. **<http://www.google.co.in/inputtools/windows>.