iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 8 **Issue:** XI **Month of publication:** November 2020

**DOI:** https://doi.org/10.22214/ijraset.2020.32273

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 8 Issue XI Nov 2020- Available at www.ijraset.com*

# A Scalable Method for Conversion of Typed Text into Handwritten Text

Suyash Mahesh Bahrani[1], Rohit Manish Punjabi[2]
[1, 2]*Department of Information Technology, Thadomal Shahani Engineering College, Mumbai, India*

***Abstract:*** ***There are various things we humans have in common. However, other things are unique to every individual - DNA, fingerprints, etc. Handwriting is one other such thing unique to each individual, which the recent studies on Handwriting analysis have already proved. As computerization is becoming more conspicuous nowadays, text transformation to handwritten text is learning significance in several fields, e.g., making notes in your handwriting is simpler within the learning cycle than learning through a text document. There are numerous situations where assistants must compose a book in their handwriting as an administrative work measure. Furthermore, letting a large scale computational systems do all the analysis and convert text to handwriting reduced much of the burden. Did you know that an average person's handwriting speed is 13 words per minute, and their typing speed is three times faster? An average person will find handwritten work experience to be awful if there is a lot of work to be done. Too much writing takes too long. It causes you to focus less on what you're writing, and instead, you paint letters on a page perfunctorily. So what to try and do when faced with unreasonably substantial work? Pine away at your scratchpad for quite a long time, composing stuff you could write in a word processor in a small amount of time? However, how would a computer convert digital text to the handwriting of an individual? Since each individual has its way of presenting his/her ideas on paper, there is a certain level of complexity involved in this subject. An overview of some methodologies and machine learning algorithms are presented here.***

***Keywords:*** ***Typed text feature extraction, detection, conversion, Handwritten font, Optical Character Recognition (OCR), Handwriting style.***

## I. INTRODUCTION

All the fashionable inventions within the computer and communication technologies like word processors, fax machines, and e-mail have an impression on handwriting. The worldwide adoption of digital messaging has given handwritten communication a special status as the approach for important, personally expressive messages. However, what can an author do if, take as an example of, suffering a stroke changes their handwriting, impacting their style and hurting the legibility of their messages? Our proposed system is a new alternative that gives the author the opportunity to render and print new text in their original pen on paper handwriting style. As a single input to our system, an end user annotates an author's handwritten text sample and follows a printer calibration processing. The user can then ask our system to turn any new text they like into handwriting that looks like the original author wrote it. These in-variations have prompted the tweaking and rethinking of the art of handwriting. Despite these modern marvels, a pen and paper are much more convenient than a keyboard or a mouse. Realistic handwritten text has many other pragmatic applications. Gifts ordered from online merchants and flower shops can include personal messages, but they're rendered in an impersonal font. Handwritten text completes the personalization because the person giving the gift could use their actual handwriting. This could also include the option of using a celebrity's handwriting on different products as a limited edition. There are similar arguments that apply to the likelihood of handwritten mail merge, e.g., for Christmas cards. Various Comics books consist of handwritten lettering; synthesis allows the style to be maintained during localization to a foreign language. Moreover, the handwritten text allows for numerous creative uses, e.g personalized books—sensitive materials like signature on credit cards. The transcribed information base contains 26 preparing pictures, 26 approval pictures, and 26 test pictures. Text to handwriting is the technique that involves image capture, enhancement, and generating the text in your handwriting.

## II. LITERATURE REVIEW

Abhishek Bal and Rajib Saha1 in 2006 [1], proposed a skew detection and normalization method that uses the orthogonal projection. The proposed method was tested on the sample database. The proposed method can detect the precise skew angle and is also ready to normalize the skew angle. The experimental result shows that the proposed system achieves higher and better accuracy for almost all types of skew angles.

*A. Text Synthesis*

According to [Wei et al., 2009] [2] handwriting synthesis bears inumerable similarities to plain text synthesis. In both cases, our goal is to generate plausible new images from existing examples. [Simoncelli., 2000: Lewis., 1984 and Portilla] [3] suggested early texture synthesis approaches were parametric. Realistic synthesis of a much more extensive range of textures became possible with the advent of non-parametric methods that sampled image patches from the input example [Efros and Freeman 2001] [4]. This was a significant shift but came at the expense of sacrificing artistic control via parameter adjustment. [Kwatra et al. 2003] [2] proposed several text synthesis approaches which can be explicitly supervised. This is simulated by restricting flowers at the top position of the canvas to come from scaled down versions of the input. Then the user takes the indicated line continuity into account when inpainting which is given by [Sun et alf. 2005] [5] . Our model is additionally a sort of structured texture synthesis, where the text being rendered is parameterized by what the user types in, but the design is nonparametrically sampled from a tagged image of the first pen-on paper handwriting sample. From a user's perspective, this is ideal.

*B. General Handwriting*

We collect data offline. Handwritten text features are analyzed at several scales, from slices of pixel to timing information, if it is available. With features extracted, a classifier then recognizes either entire words or individual characters. We use a pixel-slice type approach to recognize individual characters. Writer adaptation [Connell and Jain 2002] [6] involves adjusting the pre-trained model to match a specific writer and to improve recognition accuracy. [Shilman et al. 2006] adjusted their model as the user corrected errors in the output. Synthetic text generation has proven very beneficial as well as successful for enhancement and improving printed word recognition in the wild proposed by [Parizi et al. 2014] [7].

[Kautz and Campbell.,2014] [8] suggests that methods for automatically interpolating fonts with similar topology exist but suffer the same limitations. Accuracy of replicating an individual's handwritten text with current font engines is prohibitively difficult, to tackle this we generate textures directly. Guyon [1996] [9] acquired common two and three-letter sequences with a tablet to overcome some of the font-like approach's limitations. However, their results appear unnatural for joined up handwriting, as the two and three-letter sequences are joined up. However, they are, to the best of our knowledge, the only other paper to consider rendering pen ink and use pressure and angle of travel to drive a calligraphy-like model. In our system, we sample the texture of the original pen from the paper which is in the form of sample images, instead of rendering ink like curves, with the option to substitute other handwritten text fonts. Garner [2005] [10] demonstrated that including a handwritten note with a survey can increase the response rate from 33% to 70% – more than double. Our overarching contribution is designing a system that renders new text in a specific person's pen-on-paper handwriting style.

## III.PROPOSED WORK

If the sample of an individual's handwriting is provided, the goal is to render realistic looking handwriting within the same style, with user-specified words. There are two phases: First, the author's sample, Fig.1, is captured and analyzed. Second, the model delivers every client gave target text as a "handwritten" output, Fig.1. Finally, additional post-processing may occur, counting on the utilization case. The input to the system is a sample of an author's digitized writing, e.g., a scanned document written text. Before synthesis can occur, the sample must be analyzed. Tagging is semi-automated, consists of extracting the text's path and then segmenting it into a labeled handwritten font. Once an author's writing sample is analyzed, it is often render new text as "handwriting."

For each individual, a sample of their handwriting or handwritten text is required to drive our model. The captured handwriting must be representative of the author's style. There are two scenarios: Uploading the document or Directly paste the text into the given text area box. We optimize the source text given by the author to obtain sufficient samples with the least effort. Our aim is to construct a source text that includes the most commonly used letter combinations in this collaborative scenario but is short enough to burden the author unduly. Obviously, for a short text, the user would not use a computer at all. The source text consists of complete sentences written by the author on the back of the single-sided ruled paper or a blank page, in the pen of their choosing. They can see the ruled lines through the paper. It is necessary to perform several document analysis operations before converting text into handwriting. Some of the everyday operations performed before conversion are:

1) *Thresholding:* The task of converting the file or text into a gray-scale image and parsing the arguments.
2) *Remove Noise Present in the Text:* The system checks if there is noise in the source text and applies median blurring to remove this noise, causing errors in the conversion.
3) *Optical Character Recognition:* The converted gray-scale images are then processed through the technique of OCR. The OCR tool pytesseract reads the text embedded in the images or text area and prints the recognized text instead of writing it to a file.

After these operations are carried out, we get a handwritten document that consists of our handwritten style.

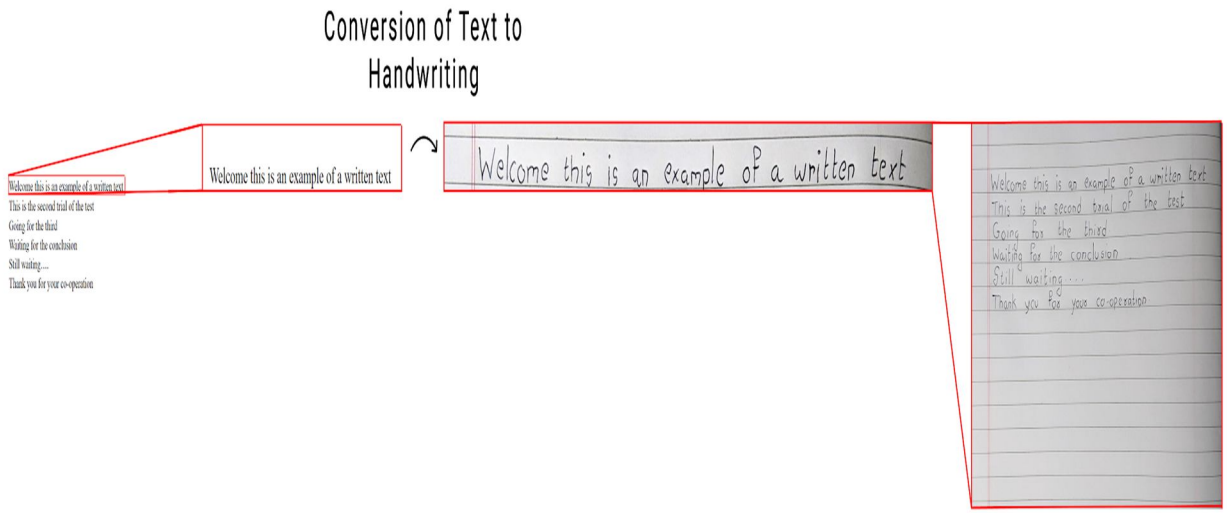



Fig. 1 Example of the system processing a typed text document into the user's handwritten text

## IV. EXPERIMENTAL RESULTS

The system encapsulating our algorithm was validated experimentally to measure the realism and style-similarity of handwritten text and demonstrate the typed text to the handwritten font. We have presented a text conversion system that can convincingly replicate a specific author's handwriting for the given text in the document. It is semi-supervised, as a user selects the target text, which is the interaction required in our use cases. The precise text conversion required to create clear handwriting is then generated to match the target text without further user input (unsupervised). We used our dataset for quantitative evaluation, to which we added samples consisting of different handwritten fonts for qualitative results. At first, noise removal techniques are applied to the handwriting text document if noise is present. Picture quality is improved by eliminating noise from the handwritten text report. The thresholding technique is applied on a noiseless grayscale handwriting image to convert the image into a grayscale image. Then the proposed method is applied for the conversion. During the recognition process, handwritten documents should free from the unbalanced skew angle for better recognition. Here the proposed method calculates the threshold value to differentiate between inter-word and intra-word gaps.

The input typed in Figure 2 is a Times New Roman typed text in a document. Figure 3 shows sentences synthesized in the style of the author handwriting style. We have only explored the English alphabet so far. There are two methods for the conversion of text to handwriting depending upon the source text.

*1)* The source text is given in a form of document and the final output is given in an final_output.pdf form by the processed system.

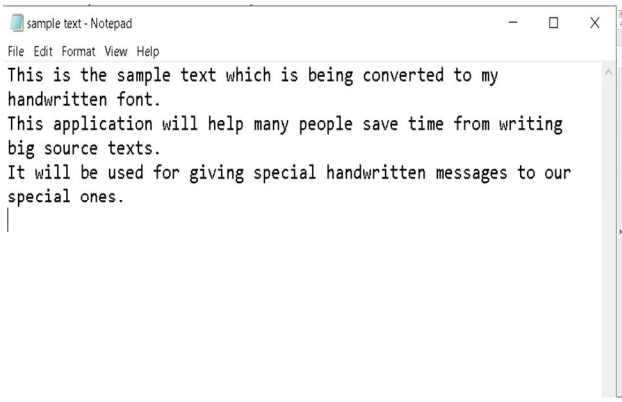



Fig. 2 Example of Sample Text in a document

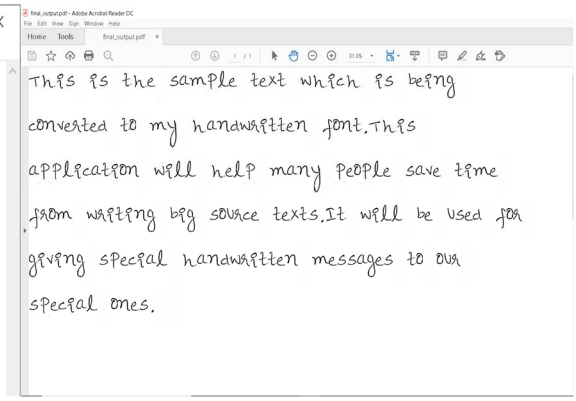



Fig. 3 Example of the output in user's handwritten font

*2)* The content to be converted into preferred handwriting is directly allowed to paste on the local host website in the text area section. Figure 4 below shows the representation of the model.

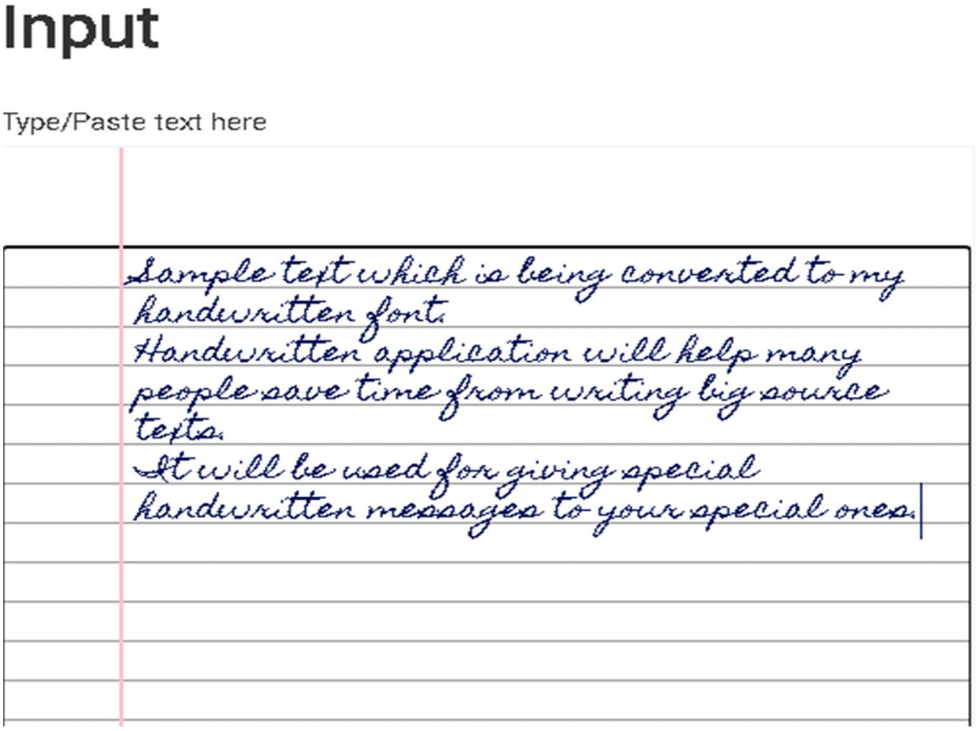


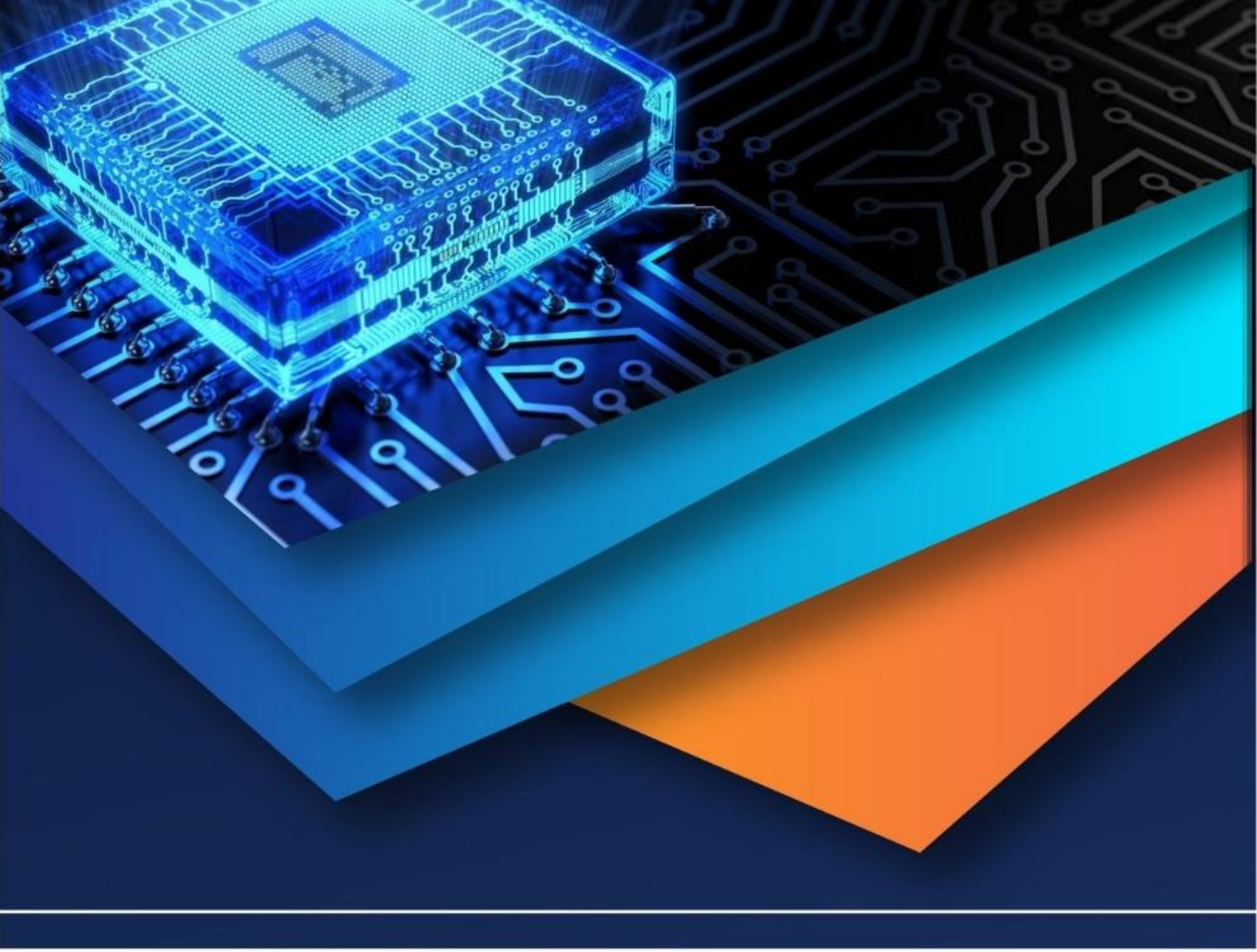

Fig. 4 Example of directly pasting the content which is converted into individuals handwriting font

## V. CONCLUSION AND FUTURE WORKS

This paper proposed a text to handwritten conversion document analysis using skew recognition, writing pressure detection, and text segmentation of cursive typed text documents that can convert into the user handwritten style. The proposed method has been applied to more than 500 text images and sample handwriting images collected from different writers on different backgrounds. Utilizing the proposed technique, 93.47% lines, and 90.41% words are accurately segmented from the dataset. The proposed work additionally standardizes 95% lines and words impeccably with negligible blunder rates. The proposed technique could be applied to different dialects' composing styles with pretty much a similar exactness. Future Work has given an outsized enough database of samples, the coverage issue may be settled by discovering writers with comparable handwriting. Our optimization strategy assumes the output goes to a blank canvas. Instead, document-repair may be possible by adding another section for the ruled pages system such that the rendered output must be on the given lines of the page. It can include more handwriting features with the proposed method like character recognition and other personality traits deciding to supersede the shortcoming and some steady factors and get a more powerful framework. As a whole, future objective is to develop a tool for behavioral analysis which can anticipate character attributes with the guide of computers without human association.

## REFERENCES

[1] Abhishek Bal, Rajib Saha., An Efficient Method for Skew Normalization of Handwriting Image., Chandigarh;2016.pp. 222-228
[2] Wei, L.-Y., Lefebvre, S., Kwatra, V., and Turk, G. State of the art in example-based texture synthesis. Eurographics.2009
[3] Portilla, J. and Simoncelli, E. P.. A parametric texture model based on joint statistics of complex wavelet coefficients. 2000
[4] Efros, and Freeman, W. T.. Image quilting for texture synthesis and transfer., SIGGRAPH: 2001. 28, 341–346.
[5] Sun, J., Yuan,. Image completion with structure propagation. SIGGRAPH: 2005
[6] Connell, and JAIN,. Writer adaptation for online handwriting recognition. PAMI: 2002., 24(3), 329–346
[7] Parizi, VEDALDI, ZISSERMAN, and FELZENSZWALB, P. Automatic discovery and optimization of parts for image classification. arXiv preprint, arXiv:2014
[8] Campbell, and Kautz,. Learning a manifold of fonts. SIGGRAPH:. 2014
[9] Guyon, I. Handwriting synthesis from handwritten texts. Frontiers of Handwriting Recognition, 1996., 309–312.
[10] Garner, R. Post-it note persuasion: A sticky influence. J. Consumer Psychology . 2005., 15, 230–237.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY