Robust Algorithm for Telugu Word Image Retrieval and Recognition

Kesana Mohana Lakshmi, Tummala Ranga Babu

1Ph. D. Scholar, Department of Electronics and Communication Engineering
Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
2RVR & JC College of Engineering, Guntur, Andhra Pradesh, India.

Corresponding Author: K. Mohana Lakshmi
Email: trbaburvr@gmail.com

Abstract

The most challenging task is searching Telugu script from the database because of difficulty in differentiating the Characteristics of the Telugu word or scripts. In this, we introduced robust approach for Telugu script retrieval using transformation-based methodology. Non-subsampled contourlet transform (NSCT) is utilized for texture classification which will function based on Non-subsampled pyramid filter bank (NSPFB) and Non-subsampled directional filter bank (NSDFB). Spatial dependence matrix is utilized to extract the texture features. In addition, image statistics is computed to enhance the retrieval performance further. Finally, hamming similarity metric is calculated which calculates the distance between trained and test word templates, which an effective distance metric over conventional Euclidean distance. In order to test, missing segment, noisy, corrupted and occlusion effected words are used as an input and taken into consideration multi conjunct vowel consonant clustered word images for showing the robustness of presented algorithm. In the substantial simulation analysis gives the presented technique finds most similar word images from database although if it is under testing conditions. Our presented scheme has superior performance compared to the traditional approaches described in the literature with respect to mean Average Precision (mAP) and mean Average Recall (mAR).

Keywords: Telugu script, texture features, statistical properties, non-subsampled contourlet transform, statistical parameters, feature vector and hamming distance metric.

I. Introduction

The process of automatic script retrieval or recognition is made from a several stages like acquiring an image from sensors, registering of images, pre-processing, extracting objects from those images, standardization or windowing, extracting
images, arrangement and post-processing. In recent years, retrieval of image document from the information of query word has emerged as a vital research field. This is also a successful alternative development for the applications of recognition-based handwritten and printed documents in most of the complex scripts retrieval system. Technically, these approaches retrieve the relevant words from the database by utilizing the extracted features of query word, which will be extracted based on various methodologies with similarity measurement. Generally, these retrieval systems performances depend on how effective the features are obtained from the word image and the similarity measuring approach utilized to match with data base word images. There are plenty of research works presented in the literature of English word images. Investigation of south Indian languages like Telugu has not been presented in much research papers in the field of word retrieval. For structuring of bag-of-words (BoW) representation, processed features using the scale-invariant feature transform (SIFT) [V] at intrigue focuses are familiar features. During content retrieval, every archive represents an irregular positioning the non-unmistakable words available in the record, showing respect to the language structure along with character arrangement. The formal Report is regarded with the help of repetition of events (histogram) of the words in the vocabulary. The histograms are later utilized to do classification of record and retrieval.

In word images, spatial sequential information of the word[XXXIII], which separates words as flat bearings hence. To represent the arrangement, word images are segregated vertically into three sections and recorded later. To deal with the fluctuation in the word style, this kind of words images are coordinated utilizing dynamic time wrapping [XXXVI, XVI], 10 Million pages were considered in a list based on arrangement hashing. In that work, pages are made into a list, retrieval is carried out in sub-seconds. The method highlights on pages as inquiry and variations of them in imaging, particularly between the camera catches. The prominence is on accurate retrieval at word level. Delayed attempts [XXX], [XII] on good record retrieval use BoVW to represent and coordinate word images. BoVW representation, and a reversed ordering plan, important documents can be retrieved from a Million documents in the span of sub-seconds [XXX].The Vector Quantization approach selects the best representation of visual word, ignoring the value of various hopefuls. Codeword believability specifies the process of selecting a codeword not having a reasonable hopeful in the vocabulary. The codebook method allocates the best suitable codeword, even that codeword will not be an authorized Specimen. In order to fix the restriction, authors in [XII], [XI] described a soft assignment coding method, where all the visual words gets allocated having a local feature depending on the position. This proposition of word image retrieval in [XXXI] briefs the BoVW method using SIFT descriptor where several languages word images are being used to check the efficiency of SIFT + BoVW in terms of mAP and mAR. In recent
times[XXVII], the author has proposed a coordination and hidden markov model (HMM) depending on Telugu word identifying with high performance over traditional word retrieval systems. In [XVII], author has explained an algorithm depending upon speeded up robust features (SURF) taking the support of BoVW to increase the performance over BoVW and SIFT + BoVW. In [XVIII], authors presented a novel hybrid methodology to retrieve and recognize Telugu word, that uses the wrenching out of texture properties features by bringing iterative partitioned clustering (IPC) to use to classify the Telugu words. On the other hand, the statistical feature removal and likelihood to match performance is additionally made better which calculates the likelihood among templates of trained and test word images. Apart from that, characters in south Indian scripts as in Telugu those which have more number of objects that makes it very difficult to use high level feature extraction methods. Specifically, missing segment Telugu word images have not been considered in earlier works.

Therefore, we introduced an efficient approach for Telugu script recognition and retrieval using transformation-based methodology. Non-sub sampled contourlet transform (NSCT) is utilized for texture classification which will function based on non-subsampled pyramid filter bank (NSPFB) and non-subsampled directional filter bank (NSDFB) with Spatial dependence matrix which is computed to extract the texture features. In addition, image statistics is computed to enhance the retrieval performance further. Finally, hamming similarity is calculated which calculates the likeness among trained and test word images. We use missing segment, noisy, corrupted and occlusion effected word images as a query input in order to test, also taken into consideration multi conjunct vowel consonant clustered word images for showing the robustness of presented algorithm. The new contributions of this work is outlined as below:

New use of multi scale decomposition and directionality using NSPFB and NSDFB respectively for classification of Telugu words. In addition, gray scale matrix (GSM) for extracting the texture from input words and image statistics is utilized for further enhancement of retrieval system.

Completely novel framework on Telugu word retrieval and recognition procedure is presented using a transformation-based proposition various kinds of Telugu word like missing segment, occlusion affected, noisy and random distortions.

In order to withstand the disadvantages of BoVW [XXX], SIFT + BoVW [XXXI], algorithm explained in [XXVII], SURF + BoVW [XVII], hash coded hamming (HCH) technique and authors presented work in [XVIII], our present work aims at performing multiscale decomposition and directionality given in section 3, extraction of texture by finding the spatial dependence matrix there by computing texture properties like contrast, correlation, energy and homogeneity, for grouping, we used
NSCT framework. We have calculated the image statistics like mean, standard deviation, variance, smoothness, skewness, kurtosis and Root mean square error (RMSE) followed by hamming similarity measurement for Telugu word image retrieval/recognition system.

The remaining paper is structured as in the given order: Section II describes in detail the relevant work carried out in the field of word image retrieval applications from the past years. Background work and motivation of NSCT framework is given in section III. The presented framework is explained in Section IV. Section V describes the results and comparative calculations against several works explained in the literature. In Section VI gives conclusion continued by references.

II. Relevant work

Author in [XXVI] explained identifying a word perceptive content; Gradient Local Auto-Correlation (GLAC) feature is very hearty and successful which are identified that to recognize or distinguish the Telugu content, the inclination feature is most suitable, feasible that the customary texture features. In [XXIV], three various features which are Gabor, Zernike moments and slope with 400 measurements for word-by word content differentiating proof and categorization has been done by using the support vector machine (SVM). The creators have made clear that the required pre-processing methods are needed to overcome the issues of the info or source. Format coordinating plan has been utilized to acknowledge word in [XXIX] with the feature ID as Gradient Angular Features (GAF) tried on 760 words using six distinctive contents. Observing all the investigated papers discussed also in the writing exposed that the new way of expressing the structure of content to check agreement to perceive or recognize the word-content. Danish et. al. [VI] use data layout coordinating technique to identify manually written, machine printed, sort composed English characters and numerals obtained a trueness of about 94.50% for usual typewritten text styles, 88% for unclear kind composed textual styles, 98% for numerals and 75% in case there needs to be happening of unclear kind composed text styles. Nikhil et. al. [XXVII] joined the layout for multi textual style styles and dimensions of English content and achieved trueness of almost 90%. Mo Wenying et. al. [XXI] joined the layout coordinating evaluation by changing related weighted coordinating degree. This calculation results in a larger coordinating rate and including deceptive identification given by traditional figuring method including trueness up to 100%.

Jatin et. al. [XIV] used structure coordinating technique to sort composed English characters and categorize using neural system classifiers. Soumendu et. al. [XXXIV] presented a computation to Japanese character recognition using the main point of gravity features and Euclidean differentiation features including character with least Euclidean differentiation is the feature used for character recognition. Mahabubar
et.al [XXII] presented a method to recognize Bangla written by hand characters using the convolution neural systems. As a mandate, abnormal state feature wrenching out is used to identify and recognize the state of a exception in the image. Shape illustration explains feature the position, introduction and size of a question in image. The representation of a shape depends on the preconditions of the application. The features that doesn’t vary of shape are practically taken into consideration to create persistent quality and power of the framework. The articles in image that search the shape data needs windowing it to a size, scale and adequate amount of brightness to differentiate the shape from its back-ground force. Various options in similar way are there to play out the coordination among the test format and produced layouts in the image[XIII].

During the recent times, author in [XXVII] described a HMM dependent correlation technique to recover the Telugu word image in [XXIV], template matching method was launched to recover the word. SURF descriptor-based k-means clustering method was calculated and examined in [XV] for which the author assumed noisy word inputs to recover the applicable word images from the database. Although, absence of efficiency, higher difficulty, lower accurateness with large data bases are the restrictions of the word image retrieval algorithms explained in the literature. In addition to that, Telugu word image retrieval is tough and demanding job because of which every word image contains its own structure including single and multi-conjunct vowel consonant cluster. We require ultimate productive features that explains the applicable data of query word images to retrieve those kind of words, this inspires us to execute a high-level feature wrenching out along with a high-level clustering algorithm that has statistical computations for Telugu word image retrieval method.

III. Background: Non-subsampled Contourlet Transform

Contourlet transform (CT) is approach of two-dimensional transform [XIX] that is utilized to represent an image with several properties like multi-resolution, anisotropy, localization, critical sampling and directionality. Multi-scaling and directionality are the basis of CT. Ideally, few coefficients extracted from the word image by utilizing CT can represent the image contours effectively. Human visual system (HVS) and curvelet transform are the motivated concepts of CT. The curvelet transform [VII], which can grasp the smoothness of image contours with unlike stretched shapes and in mixture of directions. However, the curvelet transform was implemented in continuous domain only and the directionality will be different on rectangular grid. Hence, CT is developed in discrete domain with multi-resolution and multi directionality. CT is an integration of Laplacian pyramid (LP) and directional filter banks (DFB) which are double filter bank structures and the same also called as a pyramid directional filter bank (PDFB). Only one band-pass image will be obtained.
by the utilization of LP decomposition, which avoids scrambling of frequency. DFB suits for higher frequencies only since it leaks the low frequency at directional sub bands.

However, the CT is not without flaws since it is a shift variant due to the up and down samplings presence in the filter bank structures. To regain the CT multiscale and directional properties, non-subsampled filter bank (NSFB) versions of LP and DFB structures is replaced with the conventional filter banks.

IV Proposed methodology

Algorithm 1 explains the overall procedure of presented hybrid method for retrieval of the Telugu words

Algorithm 1: NSCT with image statistics-based algorithm

Step 1: Load database word images and chose a query word.
Step 2: Read the selected query word and apply gray scale conversion.
Step 3: Later, extract the texture features by calculating the spatial dependence matrix.
Step 4: Computing characteristics like contrast, homogeneity, correlation and energy from the spatial dependence matrix.
Step 5: Apply NSCT frame work (discussed in section 3) to get multi-scale decomposition and directionality to extract the lowpass directional sub band coefficients and bandpass directional sub band coefficients from the query word.
Step 6: Compute the word image statistics from the obtained low pass directional sub band coefficients for enhancing the retrieval system performance further.
Step 7: Find the feature vector from the computed texture feature properties, low pass, band pass directional sub band coefficients and image statistics.
Step 8: Compute the feature vector of query word and data base images and denote it as Fv1 and Fv2 respectively.
Step 9: Measure the similarity between the Fv1 and Fv2 using the hamming distance metric which an efficient distance metric over conventional Euclidean measurement.
Step 10: Finally, disclose the retrieved and recognized Telugu word images.

Spatial dependence matrix

It is known as the texture investigated method with statistics in which the pixels spatial relationship is considered in a range of greyscale. Texture characterization of a source image is done by computing how frequently pixel pairs with specific intensity values and the occurrence of specified spatial relationship in an image. Then after, computing the extracted measure of statistics from this obtained SDM matrix provides texture information of word image.
NSCT frame work

The structure of NSCT contains filter banks in which the 2-D frequency plane split into number of sub bands as shown in Fig. 3. Hence, the NSCT can be divided into two sections, in which the property of multistage and multi directionality is obtained at a time. The two parts named as non-subsampled pyramid filter bank (NSPFB) and non-subsampling directional filter bank (NSDFB).

Non-subsampling Pyramid Filter Bank (NSPFB): A shift-invariant filtering process is used to get the property of multi scale that achieves a decomposition of sub band like the Laplacian pyramid, which can be acquired by a two-channel non-subsampling 2-D filter bank.

Non-subsampling Directional Filter Bank (NSDFB): It is designed by merging two-channel filter banks which are sampled critically and operations of resampling and the outcome is in a tree-structured filter bank that separates the 2-D frequency plane into the directional wedges. This is utilized to obtain a directional expansion that is shift-invariant. It is implemented by removing the up and down samplers in the DFB, which is performed by triggering off up and down samplers in every two-channel filter bank in the structure of DFB tree and up sampling the filters correspondingly. On account of this gives a tree composed of two-channel NSFBs.

Fig. 1. Layout of NSCT frame work with the integration of NSPFB and NSDFB
Fig. 2 demonstrates that the construction of integration of the NSPFB and NSDFB to implement NSCT [1]. Care must be taken in NSCT construction when DFB is applied to the coarser scales of the LP. Due to the tree-structure nature of NSDFB, the response of directionality at the frequencies of lower and upper tolerates from the effect of aliasing, that could be a serious issue in later pyramid stages, where the DFB pass band region is labelled as “Good” or “Bad.” Hence, one can view that in coarser scales, the high-pass channel in consideration is filtered with the DFB pass band bad portion, which leads to the effect of extreme aliasing and in few noticed cases a substantial loss of directional resolution. This problem gets overcome by carefully up sampling the NSDFB filters. Representing the \( k^{th} \) directional filter by \( \mathcal{H}_k(z) \). Later for higher scales, we consider \( \mathcal{H}_k(z^{2^m}) \) for \( \mathcal{H}_k(z) \) where \( m \) is chosen to make sure that the good part of the response coincides with the pyramid pass band. Note that this modification. However, authors don’t require the reconstruction section in this proposal as NSCT frame work is utilized only for extracting the coefficients in multiscale and multi directional which is not possible by the conventional feature extraction approaches.

Filtering with the up sampled filter say not improve computational complication. Particularly, in a considered sampling matrix \( S \) and a 2-D filter \( H(z) \), to get the output \( y[n] \) which results in filtering \( x[n] \) with \( H(z^S) \), we utilize the convolution formula.
\[ y[n] = \sum_{k \in \text{supp}(\mathcal{A})} h[k] x[n - S k] \]

Theàtrous filtering algorithm described in [XX]. Hence, every filter presented in the NSDFB tree has the similar complication as in building-block fan NSFB. In the same way, NSP each filtering stage has the likewise complication as that obtained in the first stage.

Therefore, NSCT complication is prescribed by the difficulty in building-block NSFBs. In case every NSFB in NSP and NSDFB structures require \( L \) operations per output sample, then for a word image of \( N \) pixels the NSCT needs around \( BNL \) operations in which the number of sub bands are denoted by \( B \).

**Image statistics**

Statistics is enquiring of the gathering, connection, examination, and execution of information. It controls all parts of this, along with the organizing information collected as long as the plan of studies and tests. The various statistical estimates [IX] are mean (\( \mu \)), variance (\( V \)), standard deviations (STD), root mean square error (RMSE), Kurtosis (K) and Skewness (S) estimates are used as a part of large variation of logical and social research, which also includes: biostatistics, computational science, computational human science, organize science, sociology, humanism, social research and in this way. In this, unique kind of statistical measure is used concerning image processing and copied them

**Mean (\( \mu \))**

\[
\text{Mean}(\mu) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y)
\]

In which \( I = \) input query word
\( x \) and \( y \) = space co-ordinates
\( M \) and \( N \) represent rows and columns number respectively

**Variance (\( V \))**

Variance is an estimate of how wide a number positioning is out stretched. This is one among very few descriptors of a similarity appropriation, illustrating how wide the numbers lie from the mean (desired favor). Particularly, the distinctness is one of the snapshots of a circulation. During such particular situation, it shapes some part of a steady way to handle identifying similarity disseminations. Whereas other similar techniques are being designed, that are in view of minutes which are suitable concerning scientific and calculative uniqueness. Scientifically deviation can be expressed as

\[
\text{Variance}(V) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left( I(x, y) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \right)^2
\]
Where $I$ is an input word image
$x$ and $y$ are spatial co-ordinates
$M, N$ represent rows and columns number respectively

**Standard Deviation (STD)**

This is widely used to estimate the changeability or miscellaneous diversity used like portion of statistics. When the image processing is considered it shows how much diversity or "scattering" exists from the mean or desired favor. A low standard deviation explains that the information guides upslope in the direction of near the mean, on the other hand exclusive requirement deviation explains that the data needs attention to extend finished an extensive scope of qualities.

$$STD = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (I(x,y) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y))^2}$$ (4)

**Kurtosis (K)**

Kurtosis values are represented with unified measurement of noise and resolution. Maximum kurtosis values indicate that lower noise and resolution.

The kurtosis is defined as

$$Kurtosis(K) = \frac{E(x-\mu)^4}{\sigma^4}$$ (5)

Where $\mu$ is the mean of $x$, $\sigma$ is the standard deviation of $x$ and $E(t)$ is the expected value of the quantity $t$

**Skewness (S)**

Skewness is an asymmetric measurement of the data of the sample mean. In case the value of skewness is under negative range, then the data is distributed more towards the left of the mean instead of right. If it has positive values, then the data spread out more to the right. For any perfect distribution that is symmetric, the value of skewness is zero.

The skewness of a distribution is expressed as

$$Skewness(S) = \frac{E(x-\mu)^3}{\sigma^3}$$ (6)

In which $\mu$ is the mean of $x$, $\sigma$ is the standard deviation of $x$ and $E(t)$ is the desired value of quantity $t$
Smoothness

It is a measured property by the number of derivatives it has that are continuous. Smoothness of an image is defined as

\[ \text{Smoothness} = 1 - \left( \frac{1}{1 + \sum_{k=1}^{m} I(k)} \right) \]  

(7)

Where \( I \) represent an image with \( k = 1, 2, 3, ..., m \)

Hamming distance metric

To find the distances between query word image \( I_q \) along with retrieved word image \( I_r \), a metric should be defined. There is a requirement of measurement method to explain how the query and retrieved word images are alike (bit per bit). Hence, there should be a likeness measure in which the distance value can be the number of similar bits in the given Telugu word image. The following table describes likeness truth table for the distance that needs to be defined.

Taking into consideration the nth bit of \( I_q \) and \( I_r \), along with the distance between those two is denoted as \( \Delta \) then truth table of similarity is as follows:

Table 1. Truth table of similarity measurement using hamming distance metric

| \( I_q[n] \) | \( I_r[n] \) | \( \Delta(I_q[n], I_r[n]) \) | Similarity |
|-------------|-------------|---------------------------|------------|
| 0           | 0           | 0                         | relevant   |
| 0           | 1           | 1                         | Irrelevant |
| 1           | 0           | 1                         | Irrelevant |
| 1           | 1           | 0                         | Relevant   |

V. Results

The performance assessment of presented Telugu word image retrieval system is obtained by the consideration of Telugu word images shown in fig. 4 and environment of MATLAB tool. Details are provided in the Table 2. As shown in fig. 4, multiple query words are generated with different sort of distortions. Based on these query input word images, the performance evaluation of presented system is done with comparison to the conventional algorithms to retrieve are given in this literature. Selection of query word is done in such a way that it should occur multiple times in given database.
Figure 3: Test query word images used for performing the retrieval process

Table 2. Book used for experiment

| Book   | #Pages | #Words |
|--------|--------|--------|
| Telugu | 425    | 7425   |

Fig 5. Retrieved images of query word image 1. (a) HCH scheme[XXIX]. (b) GSM-IPC[XXXI]. (c) NSCT with image statistics.
Initially, word image without any degradation in the quality of word image is given as a query word image to the implemented Telugu retrieval and recognition system, which is disclosed in Fig.5(a)-(c), Fig.6(a)-(c) and Fig.7(a)-(c). For all the three-word image samples, retrieved images are quite similar and have higher
precision and recall since all three methodologies compared obtained relevant word images.

![ Retrieved images of query word image 4 (noisy). (a) HCH scheme [XVII]. (b) GSM-IPC [XVIII]. (c) NSCT with image statistics ]

Example of noisy query word image is demonstrated in Fig.8(a)-(c), where Fig.8(a) presented the outcome obtained using the algorithm HCH scheme [XVII], Fig.8(b) discloses the GSM-IPC [XVIII] system and the NSCT with image statistics approach is explained in Fig. 8(c). Observing the Fig.8(a)-(c), it is concluded that the HCH scheme obtained more irrelevant word images while retrieving the Telugu word images with noisy inputs, whereas the GSM-IPC and NSCT with image statistics approaches obtained quite similar word image even in the noisy case.

For the case of corruption, Fig. 9 presented the obtained retrieved word images using HCH scheme, GSM-IPC and NSCT with image statistics approaches. Fig.9(a) shown that the retrieved Telugu words obtained using HCH scheme, it has shown that there are only two relevant word images out of four outcomes which leads to the mitigation of precision and recall performance. Retrieved Telugu word images using GSM-IPC approach for the corrupted query word image is shown in Fig. 9(b), where there is one irrelevant image.
Fig 9 Retrieved images of image 5 (corrupted). (a) HCH scheme [XVII]. (b) GSM-IPC [XVIII]. (c) NSCT with image statistics

Fig.9(c) shown that the NSCT with image statistics approach for the retrieval of Telugu word images in the case of corrupted query word image. This is also an example for the multi-conjunct vowel consonants clustered image. It has shown that there is no irrelevant image even with the corrupted query word image, which leads to the maximum precision and recall compared to the conventional Telugu word retrieval and recognition system.

Fig.10 shown the examples for the occlusion effected normal and multi-conjunct vowel consonant clustered Telugu word images as a query respectively. From these two outcomes, it has shown that the NSCT with image statistics approach has higher precision and recall performance over HCH scheme and GSM-IPC approach since there are few irrelevant word images while retrieving the Telugu word images.
Fig 10 Retrieved images of image6 (occlusion effected). (a) HCH scheme [XVII]. (b) GSM-IPC [XVIII]. (c) NSCT with image statistics.

Missing segment word image example is shown in Fig. 11 and Fig. 12 with comparison to the HCH scheme, GSM-IPC and NSCT with image statistics approaches, in which both the multi-conjunct vowel consonant clustered images are considered for showing the effectiveness and robustness of Telugu word image retrieval and recognition systems. From the obtained outcomes of these two figures, it is clear that the NSCT with image statistics is well suited for the Telugu word image retrieval of missing segment query word images.

Fig 11. Retrieved images of image7 (missing segment). (a) HCH scheme [XVII]. (b) GSM-IPC [XVIII]. (c) NSCT with image statistics.
Fig 12. Retrieved images of image8 (missing segment). (a) HCH scheme [XVII]. (b) GSM-IPC [XVIII]. (c) NSCT with image statistics.

Fig.9 and fig 10 gives the Retrieved output images of multi- conjunct vowel consonant clustered word images including effects of occlusion along with corruption. The corrupted and occlusion effected multi- conjunct vowel consonant clustered word images results in some irrelevant words with the algorithm presented in ref [XVIII] by the author and missing segment words has not been considered in the work presented by the author. We demonstrated the retrieved Telugu word images using the presented approach, which enhances the performance of retrieval system by utilizing the multi-scale decomposition and directionality with image statistics. Fig.12 depicted that the retrieved Telugu word images with the missing segment words and a query. Retrieved words of corrupted and occlusion effected query word images shown in fig.9 and fig 10. It is shown that the presented approach gives the more relevant while retrieving these sorts of query word images, which was not done in previous work [XVIII].

Table 3. mAP and mAR of different schemes on different books in dataset

| Measurement | BoVW [XII] | SIFT + BoVW [XXXI] | Ref [XXVII] | Ref [XVII] | Ref [XVIII] | Proposed |
|-------------|------------|---------------------|-------------|------------|------------|----------|
| mAP         | 0.817      | 0.853               | 0.89        | 0.731      | 0.967      | 0.998    |
| mAR         | 0.77       | 0.799               | 0.823       | 0.809      | 0.842      | 0.98     |
VI. Conclusions

Implementation of robust algorithm for Telugu word image retrieval and recognition system using NSCT framework using spatial dependence matrix and image statistics having various kinds of word images as a query is carried out in this work. It is highly concentrated to obtain the most relevant images though the query word image affected with noise, occlusions, random distortions and even that of missing segments. This paper work targets at implementing texture features extraction by computing spatial dependence matrix and utilized NSCT algorithm for multi-scale decomposition and directionality. Also computed the image statistics proceeding with hamming similarity measurement to enhance additionally in the presented retrieval system. Considerable simulation analysis impart that the presented methodology is robust and achieved a mAP of 0.998 and mAR of 0.98 gives superior performance over the conventional retrieval system.
References

I. Arthur L. da Cunha, Jianping Zhou, and Minh N. Do, “The Non-subsampled Contourlet Transform: Theory, Design and Applications”, IEEE Transaction on Image Processing, Vol. 15, No. 10, pp. 3089-3100, 2006.

II. B. Verma, M. Blumenstein, S. Kulkarni, “Recent achievements in off-line handwriting recognition systems”, School of Information Technology, Griffith University, Gold Coast Campus.

III. C. V. Jawahar and A. Kumar, “Content-level Annotation of Large Collection of Printed Document Images”, In: Proc. of International Conf. on Document Analysis and Recognition, Parana, Brazil, 2007.

IV. C. V. Jawahar, M. N. S. S. K. Pavan Kumar, S. S. Ravi Kiran, “A Bilingual OCR for Hindi-Telugu Documents and its Applications”, Centre for Visual Information Technology, International Institute of Information Technology, Hyderabad.

V. D. G. Lowe, “Distinctive Image Features from Scale-Invariant Key points,” International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.

VI. Danish Nadeem and Saleha Rizvi, “Character recognition using template matching”, Department of Computer Science, JamiaMilliaIslamia, New Delhi, 2015.

VII. E Candes and D. Donoho, "Curvelets – a surprisingly effective nonadaptive representation for objects with edges." In: A. Cohen, C. Rabut and L. Schumaker, Editors, Curves and Surface Fitting: Saint-Malo 1999, Vanderbilt University Press, Nashville, pp. 105–120, 2000.

VIII. E. Kreyszig, Advanced Engineering Mathematics, J. Willey & Sons Inc. 2011.

IX. E. Kreyszig, Advanced Engineering Mathematics, J. Willey & Sons Inc. 2011.

X. I. Z. Yalniz and R. Manmatha, “An Efficient Framework for Searching Text in Noisy Document Images”, IAPS International Workshop on Document Analysis Systems, Gold Cost, QLD, Australia, pp. 48-52, 2012.

XI. J. van Gemert, C. J. Veenman, A. W. M. Smeulders, and J.-M. Geusebroek, “Visual Word Ambiguity”, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 32, No. 7, pp.1271-1283, 2010.

XII. J. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. M. Smeulders, “Kernel Codebooks for Scene Categorization”, In: Proc. Of European Conf. on Computer Vision, Berlin, Heidelberg, pp. 696-709, 2008.

XIII. Jangala. Sasi Kiran, N. Vijaya Kumar, N. SashiPrabha and M. Kavya, “A Literature Survey on Digital Image Processing technique in character recognition of Indian languages”, International Journal of Computer Science and Information Technologies, Vol. 6, No. 3, pp. 2065-2069, 2015.
XIV. Jatin M Patil and Ashok P. Mane, “Multi Font and Size Optical Character Recognition Using Template Matching”, International Journal of Emerging Technology and Advanced Engineering, Vol. 3, No. 1, pp. 504-506, 2013.

XV. K Mohana Lakshmi and T Ranga Babu, “Searching for Telugu Script in Noisy Images using SURF Descriptors”, IEEE 6th International Conference on Advance Computing, pp: 480-483, 2016.

XVI. K. Takeda, K. Kise, and M. Iwamura, “Real-time document image retrieval for a 10 Million pages database with a memory efficient and stability improved LLAH”, International Conf. on Document Analysis and Recognition, Beijing, China, pp. 1054-1058, 2011.

XVII. K.Mohana Lakshmi, Dr.T.Ranga Babu, “A Novel Telugu Script Recognition and Retrieval Approach Based on Hash Coded Hamming”, ICCCEP(Springer LNS), 978-981-13-0211-4, 2018.

XVIII. KesanaMohana Lakshmi and TummalaRangaBabu, “A New Hybrid Algorithm for Telugu Word Retrieval and Recognition”, International Journal of Intelligent Engineering and Systems, Vol. 11, No. 4, pp.117-127, 2018.

XIX. M N Do and M Vetterli, “The contourlet transform: an efficient directional multiresolution image representation”, IEEE Transactions on Image Processing, Vol. 14, No. 12, pp. 2091-2106, 2005.

XX. M. J. Shensa, “The discrete wavelet transform: Wedding the àtrous and Mallat algorithms,” IEEE Trans. Signal Process., Vol. 40, No. 10, pp. 2464–2482, 1992.

XXI. M. Wenying and D. Zuchun, “A Digital Character Recognition Algorithm Based on the Template Weighted Match Degree”, International Journal of Smart Home, Vol.7, No. 3, pp. 53-60, 2013.

XXII. Md. Mahbubar Rahman, M. A. H. Akhand, Shahidul Islam, Pintu Chandra Shill and M. M. Hafizur Rahman, “Bangla Handwritten Character Recognition using Convolutional Neural Network”, International Journal of Image, Graphics and Signal Processing, Vol. 7, No. 8, pp. 42-49, 2015.

XXIII. N. Sharma, S. Chanda, U. Pal and M. Blumenstein, “Word-wise Script Identification from Video Frames”, In: Proc. of International Conf. on Document Analysis and Recognition, Washington, DC, USA, pp.867-871, 2013.

XXIV. N. Shobha Rani Vasudev T and Pradeep C.H. “A Performance Efficient Technique for Recognition of Telugu Script Using Template Matching”, International Journal of Image, Graphics and Signal Processing, Vol. 8, No. 3, pp.15-23, 2016.

XXV. N. Shobha Rani, T. Vasudev, “A Generic Line Elimination Methodology using Circular Masks for Printed and Handwritten Document Images”, Emerging research in computing, information, communication and applications ELSEVIER science and technology, Vol. 3, No. 1, pp. 589-594, 2014.
XXVI. N. Sharma, U. Pal, and M. Blumenstein, “A Study on Word Level Multi-script Identification from Video Frames”, In: Proc. of International Joint Conf. on Neural Networks, Beijing, China, pp. 1827-1833, 2014.

XXVII. Nagasudha D and Y Madhavi Latha, “Keyword Spotting using HMM in Printed Telugu Documents”, In: Proc. of International Conf. on Signal Processing, Communication, Power and Embedded Systems, Paralakshemundi, India, pp: 1997-2000, 2016.

XXVIII. Nikhil Rajiv Pai and Vijaykumar S. Kolkure, “Design and implementation of optical character recognition using template matching for multi fonts size”, International Journal of Research in Engineering and Technology, Vol. 4, No. 2, pp. 398-400, 2015.

XXIX. P. Shivakumara, N. Sharma, U. Pal, M. Blumenstein, and C. L. Tan, “Gradient-Angular-Features for Word-wise Video Script Identification”, In: Proc. of International Conf. on Pattern Recognition, Stockholm, Sweden, pp. 3098-3103, 2014.

XXX. R. Shekhar and C. V. Jawahar, “Word Image Retrieval Using Bag of Visual Words”, IAPS International Workshop on Document Analysis Systems, Gold Cost, QLD, Australia, pp. 297-301, 2012.

XXXI. Ravi Shekhar and C V Jawahar, “Word Image Retrieval Using Bag-of-Visual Words”, In: Proc. of IAPR International Workshop on Document Analysis Systems, Gold Cost, QLD,

XXXII. Rinki Singh, Manideep Kaur, “OCR for Telugu Script Using Back-Propagation Based Classifier”, International Journal of Information Technology and Knowledge Management, Vol. 2, No. 2, pp. 639-643, 2010.

XXXIII. S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories”, In: Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, New York, USA, 2006.

XXXIV. Soumendu Das and Sreeparna Banerjee, “An Algorithm for Japanese Character Recognition”, International Journal of Image, Graphics and Signal Processing, Vol. 7, No. 1, pp. 9-15, 2014.

XXXV. Suman V Patgar, Vasudev T, Murali S, “A system for detection of fabrication in photocopy document”, Journal of Computer Science & Information Technology, Vol. 5, No. 14, pp. 29–35, 2015.

XXXVI. T. M. Rath and R. Mannmatha, “Word spotting for historical documents”, International Journal of Document Analysis and Research, Vol. 9, No. 2-4, pp. 139-152, 2007.