Kannada Text Image Matching (KTIM) Algorithm

Mrs.Tanuja K, ISE Dept,
Acharya Institute of Technology, Bangalore – 90

Mrs. UshaKumari, ISE Dept,
Acharya Institute of Technology, Bangalore - 90
Email: ushavekappa@gmail.com

Abstract – In this paper we present an algorithm to retrieve Kannada text image from huge collection of Kannada image using the Kannada text matching algorithm based on BLPOC (Band - limited phase only correlation) function which provides the higher correlation peak and better discrimination capability. This algorithm can achieve the effective and efficient search and retrieval from a huge collection of Kannada images by matching image at word – level and not at character – level . On the given query image Kannada text matching algorithm based on BPLOC function is used to match the words and retrieves the relevant document.

Index Terms—
Morphological operation, Phase only correlation, Band - limited phase only correlation based image matching

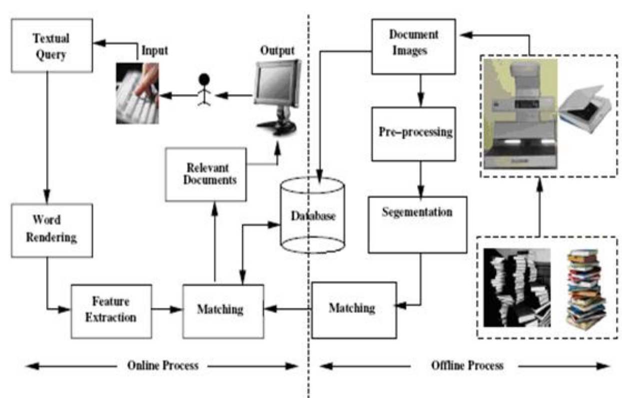
1. INTRODUCTION

A number of approaches have been proposed in recent years for efficient search and retrieval of document images. The collections of historical prints and books exist in Indian languages that need search options in images. The document images of such collections cannot be recognized accurately. There is a need for easy and efficient access to such collections. The search procedures available for text domain can be applied, if these document images are converted into textual representations using recognizers. However, it is an infeasible solution due to the unavailability of efficient and robust OCRs for Indian languages.

Addressing this problem, this paper proposes an efficient document image retrieval algorithm using Band – limited phase only correlation based image matching (BPLOC) – an image matching technique using the phase components in Inverse Fast Fourier Transformation which determines the phase angle of input image and query image that helps in matching word for the retrieval of document. This approach is faster as it does not match the image pixel by pixel.

The remainder of the paper describes our current development effort in more detail. Section III describes the architecture of the research we are developing. Section IV details the implementation procedures we are developing. Section V presents some experimental results. Section VI concludes the paper.

2. SYSTEM ARCHITECTURE



Fig(1) System Architecture

This system accepts a text query from users. The text query is initially converted to an image by word rendering, features are extracted from these images and then a search is carried out for retrieval of exact documents. Results of the search are pages from document image collections containing the retrieved words sorted based on their relevance to the query. This work mainly aims at addressing some of the issues involved in effective and efficient retrieval in document images with effective representations of the word images.

3. IMPLEMENTATION

An efficient and effective mechanism for retrieval of relevant Kannada document from a large Kannada document image collection is presented in this work. This involves three phases: first phase includes pre processing, which is preparing the source image for searching the query Kannada word by removing the noise, second phase includes generating the texted query image from the query word, and third phase includes matching of images to find the query word in the source image

In first phase, consider a source image, which is generally in form of RGB is converted to the grey scale image. Then, this grey scale image is, in turn, converted to the binary image, i This process of conversion will help to performing morphological operation. Morphological operation is performed to initiate the dilation, which helps in differentiating two words delimited by a space. Then, fill the holes to find any picture in the image and remove the big area, which might be, say, a photograph. Then, record the coordinates, height and width of each word in image document.

Second phase includes generating the query image from the query word. Convert the query image to binary image. This helps in comparison of input image with the query image.

Third phase includes matching of images to find the query word in the source image. The method used is KTIM algorithm. This phase has input as two images, source image and query image, which are converted into binary form.

4. MATCHING TECHNIQUE

In this section, the principle of phase-based image matching using the limited Phase-Only Correlation (POC) function (which is sometimes called the “phase-correlation function”)is explained . Consider two N1×N2 images, f(n1, n2) and g(n1, n2), where we assume that the index ranges are n1 = −M1 · · ·M1 (M1 > 0) and n2 = −M2 · ··M2 (M2 > 0) for mathematical simplicity, and hence N1 = 2M1 + 1 and N2 = 2M2 + 1. Let F(k1, k2) and G(k1, k2) denote the Fast Fourier Transform(FFT) of the two images. F(k1, k2) is given by

$$F(k_1,k_2)=\sum_{n1,n2} f(n_1,n_2)W_{N_1}^{k_1n_1}W_{N_2}^{k_2n_2}$$

$$= A_F(k_1,k_2)e^{j\theta_F(k_1,k_2)}$$

where k1 = −M1 · · ·M1, k2 = −M2 · · ·M2, $W_{N_1}=e^{-j\frac{2\pi}{N_1}}$, $W_{N_2}=e^{-j\frac{2\pi}{N_2}}$ and $\sum_{n_1,n_2}$ denotes $\sum_{n_1=-M_1}^{M_1}\sum_{n_2=-M_2}^{M_2}$ AF (k1, k2) is amplitude and θF (k1, k2) is phase. G(k1, k2) is defined in the same way. The cross-phase spectrum R_{FG} (*k*1, *k*2) is given by

$$R_{FG}(k_1,k_2)=\frac{F(k_1,k_2)\overline{G(k_1,k_2)}}{\left|F(k_1,k_2)\overline{G(k_1,k_2)}\right|}$$

$$= e^{j\theta(k_1,k_2)}$$

where G(k1, k2) is the complex conjugate of G(k1, k2) and θ(k1, k2) denotes the phase difference θF (k1, k2)−θG(k1, k2). The POC function rfg(n1, n2) is the Inverse Fast Fourier Transform (IFFT) of RFG(k1, k2) and is given by

$$r_{fg}(n_1,n_2)=\frac{1}{N_1N_2}\sum_{k_1,k_2} R_{FG}(k_1,k_2)W_{N_1}^{-k_1n_1}W_{N_2}^{-k_2n_2}$$

where $\sum_{k_1,k_2}$ denotes $\sum_{k_1=-M_1}^{M_1}\sum_{k_2=-M_2}^{M_2}$

When two images are similar, their POC function gives a distinct sharp peak. When two images are not similar, the peak drops significantly. The height of the peak gives a good similarity measure for image matching, and the location of the peak shows the translational displacement between the images.

Band-Limited Phase-Only Correlation (BLPOC)

In the POC-based image matching method, all the frequency components are involved. However, high frequency tends to emphasize detail information and can be prone to noise. To eliminate meaningless high frequency components, K. Ito et al. [9] proposed the Band-Limited Phase-Only Correlation (BLPOC). The BLPOC limits the range of the spectrum of the given image. Assume that the ranges of the inherent frequency band of are given by $u=-U_0,\ldots,U_0$ and $v=-V_0,\ldots V_0$, where $0<=U_0<=M_0$, $0<=V_0<=N_0$. Thus, the effective size of spectrum is given by $L_1=2U_0+1$ and $L_2=2V_0+1$. The BLPOC function is defined as

$$p_{gf}^{U_0V_0} m,n = \frac{1}{L_1L_2}\sum_{u=U_0}^{U_0}\sum_{v=V_0}^{V_0} R_{GF}(u,v)e^{j2\Pi(\frac{mu}{L_1}+\frac{nv}{L_2})}$$

where $m=-U_0,\ldots,U_0$ and $n=-V_0,\ldots,V_0$. When two images are similar, their BLPOC function gives a distinct sharp peak. Also, the translational displacement between the two images can be estimated by the correlation peak position. Experiments indicate that the BLPOC function provides a much higher discrimination capability than the original POC function

Algorithm for Kannada text Image matching:
Procedure : Word matching using BPLOC function
Input:
$f(n_1,n_2)$: the registered kannada text image
$g(n_1,n_2)$: the kannada text image need to be verified
Output:
Matching score between $f(n_1,n_2)$ and $g(n_1,n_2)$;

1. **Begin**
2. Store in advance a set of rotated images $f_\theta(n_1,n_2)$ of $f(n_1,n_2)$ over the angular range $-\Theta_{max}\le\Theta\le\Theta_{max}$ with an angle spacing 1^0;
3. Calculate the POC function $\hat{r}^{M_1M_2}_{f_\theta g}(n_1,n_2)$ between $f_\theta(n_1,n_2)$ and $g(n_1,n_2)$;
4. Calculate $\Theta=\arg\max\{S_1^{M_1M_2}[f_\Theta,g]\}$ to select the rotation-normalized image $f_\Theta(n_1,n_2)$;
5. Estimate the image displacements(τ_1 , τ_2) between $f_\Theta(n_1,n_2)$ and $g(n_1,n_2)$ from peak location of $\hat{r}^{M_1M_2}_{f_\theta g}(n_1,n_2)$;
6. Extend the size of $f_\Theta(n_1,n_2)$ and $g(n_1,n_2)$ by τ_1 and τ_2 pixels for n_1 and n_2 direction to both $f^{`}(n_1,n_2)$ and $g^{`}(n_1,n_2)$;
7. Extract the effective the text regions $f^{``}(n_1,n_2)$, $g^{``}(n_1,n_2)$ from $f^{`}(n_1,n_2)$ and $g^{`}(n_1,n_2)$;
8. Detect the inherent frequency band (K_1,K_2) from the 2D DFT of $f^{``}(n_1,n_2)$;
9. Calculate the band limited POC function(BLPOC) $\hat{r}^{K_1K_2}_{f^{''}g^{''}}(n_1,n_2)$;
10. Compute the matching score; $S_P^{K_1K_2}[f^{''},g^{''}]$
11. **end**

5. EXPERIMENTAL RESULTS

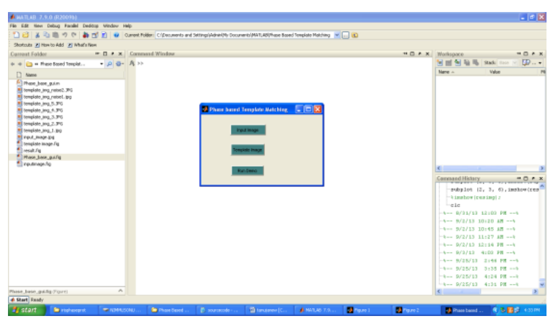


Figure (5.1) shows the input to the system

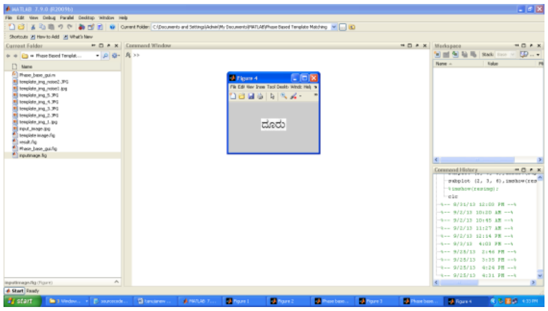


Figure (5.2) Query to the system

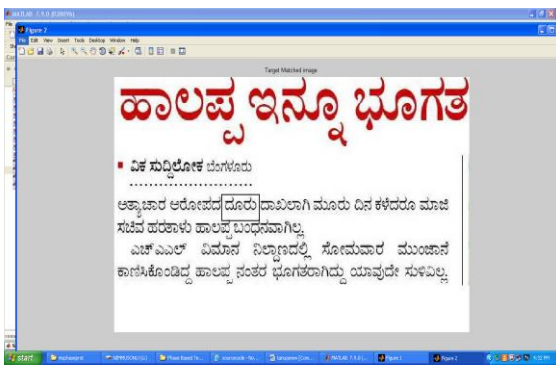


Figure (5.3) Shows the output

Since we are using database approach for the character recognition, in this approach for each character we need to have details like Character name, Character BMP image. This takes lot of space as well as lot of computation involved in recognizing the character.

6. CONCLUSION

In this paper we have presented retrieval in large document image collections. The matching technique is based on phase - based image matching, for search in large collections of document word images is applied to obtain good performance. The approaches used for word spotting so far, dynamic time warping and/or nearest neighbor search tend to be slow for large collection of books. Direct matching of pixels in images is inefficient due to the complexity of matching and thus impractical for large databases. This problem is solved by directly storing word image representations.

Some of the possible directions in which the future work can be carried out are as below. The effect of combination of different fonts in a single collection can be one possible direction for exploring the feasibility of the proposed technique and improving it. Multi-lingual document images are not handled in the proposed technique. Segmentation is not handled in this paper

7. Reference

[1] Koichi Ito, Ayumi Morita, Takafumi Aoki, Tatsuo Higuchi, Hiroshi Nakajima, and Koji Kobayashi, A Fingerprint Recognition Algorithm Using Phase-Based Image Matching for Low-Quality Fingerprints

[2] Ashwin T V 2000 *A font and size independent OCR for printed Kannada using SVM*. M E Project Report, Dept. Electrical Engg., Indian Institute of Science, Bangalore

[3] A. Balasubramanian, Million Meshesha, and C.V. Jawahar Retrieval from Document Image Collections Centre for Visual Information Technology, International Institute of Information Technology, Hyderabad - 500 032, India

[4] Ito, K., Nakajima, H., Kobayashi, K., Aoki, T., Higuchi, T.: A fingerprint matching algorithm using phase-only correlation. IEICE Trans. Fundamentals E87-A(3), 682–691 (2004)