

# Automatic Text Extraction in Video Based on the Combined Corner Metric and Laplacian Filtering Technique

Kaushik K.S<sup>1</sup>, Suresha D<sup>2</sup>

<sup>1</sup>Research scholar, Canara Engineering College, Mangalore

<sup>2</sup>Assistant Professor, Canara Engineering College, Mangalore

**Abstract**— Text in the videos and images provides very major information about the content of the videos, the video text extraction provide a major role in semantic analysis of the video, video indexing and video retrieval which is an important role in video database. We propose an efficient method for detecting, localizing and extracting the text appearing in the videos with noisy and complex background. The text region appearing in the video or an image has certain features that distinguishes it from the rest of the background, we make use of corner metric and Laplacian filtering techniques to detect the text appearing in video independent of each other and combine the results for an efficient detection and localization. Then the binarization of the localized text is done by the seed pixel determination of the text. the experimental results shows the efficiency of the proposed system.

**Index Terms**— corner metric, Laplacian filter, text detection, text localization, text binarization, text extraction, video .

## I. INTRODUCTION

Text in the videos and images provide a very major information about the content of the videos, the video text extraction provide a major role in semantic analysis of the video, video indexing and video retrieval which is an important role in video database[1] [5].

There are two types of text that can appear in the video or an image: the scene text and the artificial text, the scene text is type of text that accidentally appears in the video and doesn't provide a reliable information for video indexing and retrieval, the artificial text is a type of text that has been artificially overlaid on the video and provides a valuable information for video retrieval and indexing [7] so our focus will remain on the extraction of artificial text. Therefore the word text appearing in this paper refers to the artificial text.

Most of the video indexing research starts with the video text recognition, the process of video text recognition can be divided into four steps: text detection, text localization, text binarization and text recognition, text detection is the process of roughly differentiating the text and non-text regions of the video, text localization process involves determining the accurate boundaries of text strings, the text binarization step

involves filtering the background pixels in the text strings, these extracted text pixels are then left for recognition, the first three steps above mentioned are collectively known as text extraction. The data flow for the text extraction system is shown in the figure 1.

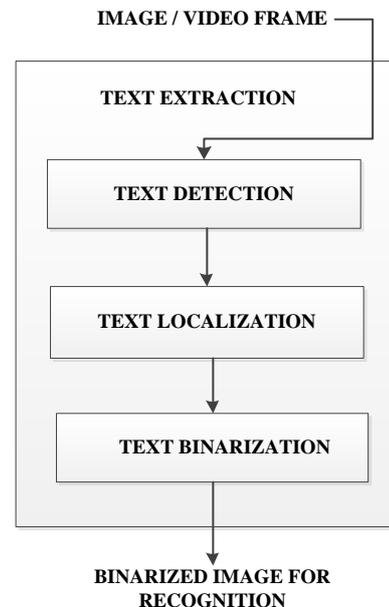


Fig. 1 The data flow for the text extraction system.

## II. LITERATURE SURVEY

The video text detection methods can mainly be classified into two categories: 1.) the connected component based method which assumes that text strings contains some unique characteristics such as uniform colors, font size and spatial alignments are satisfied, these methods usually performs color reduction and segmentation in some color space and then perform connected component analysis to detect the text regions. The main problem with this kind of method is that it is not universal for all kind of images [2], [3]. 2) Edge or texture based approach hold the assumption that text regions have specific patterns and has more edge features than the

smooth background. The main problem of this kind of method is to reduce the noise coming from the backgrounds [4], [5]. The texture based methods can still be categorized into top-down approaches and bottom down approaches. The top-down approach is based on splitting image regions into horizontal and vertical direction based on the texture, color or edge [4]. The bottom-up approach intends to find homogenous regions from some seed regions, and then the region growing technique is applied to merge the pixels belonging to the same clusters [6].

Some of the related works of text detection and localization involves the following: [7] Propose a solution for detecting a text appearing in the images and videos by making use of the corner response in the images which is obtained by implementing corner detection algorithm proposed by [25], the method doesn't detect the corner accurately and also produces a lot a noise in the image with high clarity, which adds an additional burden of removing the noise without removing the actual text present in the video or images. The method proposed by [8] computes an energy measure of a set of DCT coefficients of intra-coded blocks as a texture measure, this method obtains the seed pixels by making use of multiple iterations and by using different thresholding values for each iterations, the suitable seed pixels are calculated by decreasing the threshold values for each successive iterations, based on these pixels the text is detected and localized in the videos. [9] Propose a novel approach for efficient automated text detection in video data; firstly the potential text candidates are identified by using an edge-based multi-scale text detector. Secondly, image entropy based improvement algorithm and a Stroke Width Transform (SWT) based verification procedure are used for refining the candidate text lines, Both types of artificial text and recorded scene text can be localized reliably. [10] Propose an effective coarse-to-fine algorithm to detect text in video. Firstly, in coarse-detection part, candidate stroke pixels are detected by making use of the stroke filter, and then these pixels are connected together into a regions by making use of a fast region growing algorithm, these regions are further separated into candidate text lines by projection operation. Secondly, in fine-detection section, support vector machine (SVM) model and stroke features are employed to select the correct text regions from the candidate ones, and text regions in multi-resolution are integrated. Finally, the results are optimized significantly according to temporal correlation information. The proposed solution by [11] explores new edge features such as straightness for the elimination of non significant edges from the segmented text portion of a video frame to detect accurate boundary of the text lines in video images. The method introduces candidate text block selection from a given image, to segment the complete text portions. Based on combination of filters and edge analysis, heuristic rules are formed for identifying a candidate text block in the image. This method effectively detects and localize the text but it includes many steps to do so, which in turn reduces the performance in terms of time to detect and localize the text appearing in the video. The

solution proposed by [12] is based on invariant features, such as edge strength, edge density, and horizontal distribution. First, it applies edge detection and then filter out the verified non-text edges by using a low threshold value. Then, to both keep low-contrast text and simplify complex background of high-contrast text, a local threshold value is selected. Next, the high edge strength or high edge density is enhanced by using two text-area enhancement operators. Finally, coarse-to-fine detection locates text regions efficiently. [13] Proposes a hybrid system for text detection in video frames. The system consists of two main stages. In the first stage edge map of the image is used for the detection of text regions. In the sequel, a refinement stage uses an SVM classifier trained on features obtained by a new Local Binary Pattern based operator which results in diminishing false alarms. [14] Propose an edge based detection method which is based on an edge map produced by the Sobel operator followed by smoothing filters, morphological operations and geometrical constraints. [15] Propose a system for object detection based on Local Binary Patterns (LBP) and Cascade histogram matching. The LBP operator consists of a 3x3 kernel where the centre pixel is used as a threshold. Then the eight binarized neighbors are multiplied by the binomial weight producing an integer that represents a unique texture pattern. This proposed method is used for videotext and car detection. [16] Proposes an improved algorithm for the automatic extraction of artificially overlaid text in sports video. First step is making use of the color Histogram technique to minimize the number of video frames for the identification of key frames from video. Then, the efficient text detection is done by converting the key images into a gray scale images. Generally, the super imposed text displayed in bottom part of the image in the sports video. So, they cropped the text image regions in the gray image which contains the text information. Then canny edge detection algorithms are applied for effective text edge detection. The proposed solution [17] explores an idea of classifying low contrast and high contrast video images in order to detect accurate boundary of the text lines in video images. Were high contrast refers to sharpness while low contrast refers to dim intensity values in the video images. The classification of the text regions is done by introducing heuristic rules based on combination of filters and edge analysis. The heuristic rules are derived based on the fact that the number of canny edge components is less than the number of Sobel edge components in the case of low contrast video images, and vice versa for high contrast video images. The solution proposed by [18] handles the complexity of video text detection by combining the Fourier and statistical features in color space of the image, since it uses a large number of features in frequency and color domain, it is computationally expensive and it is limited to horizontal text detection. [19] used the homogeneity of intensity of text regions in images. Pixels with similar gray levels are merged into a group, the candidate regions are then subjected to verification using size, area, fill factor, and contrast. [20] Propose a method for text detection by combining the edge and gradient features of the

image, this method produces efficient text detection with low false positives but the detection technique works only the image where the text is overlaid in the horizontal direction. [21] Uses a combination of texture features and edge features in the wavelet transform domain with a classifier to detect text in video. [22] Locates text in images and video frames by making use of neural network classifier and image gradient features. Some methods [23], [24] assume that the text strokes have a certain contrast against the background. Therefore, those areas with dense edges are detected as text regions.

Text binarization methods can be categorized into two main groups: thresholding-based algorithms including global thresholding methods [26], which use a single threshold for the entire document, image or video generally calculated by making use of global thresholding algorithms and local thresholding methods [27], [28], [29], which assign a threshold for each small local region of the processed data. The second class sums up region-based, clustering based [30] and edge-based [31] methods.

### III. PROPOSED SOLUTION

The architecture of the proposed solution is shown in the figure 2.

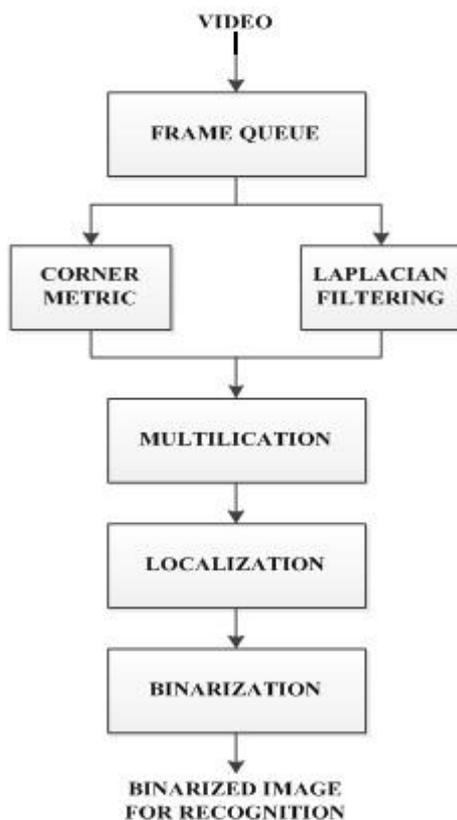


Fig. 2 The data flow for the text extraction system.

For an efficient video indexing and content based video retrieval it is necessary to detect and extract every text appearing in a video without missing any block of text, we have considered the worst case that the text appearing in the

video changes (new text appears) for each and every second but not within a second. The video frames are captured at every second and stored in a frame queue; therefore the number of frame captured and stored in the queue depends on the length of the video, for example if the length of the video is 5 seconds then 6 frames are stored in the queue, which are captured at each of the 5seconds and the frame captured at the 0<sup>th</sup> (zero) second. As shown in the figure 2 each of these frames is then processed for text detection, text localization and text binarization independent of each other.

#### A. Text Detection

The text detection part comprises of three steps: Corner Metric, Laplacian Filtering and Multiplication. Both the corner metric and Laplacian filtering technique make use of the different properties of artificial text for its detection and the independent results of both these techniques are combined to form a single result by performing the multiplication on the results of these individual results, combining the results is done because it performs a very efficient noise reduction.

1) *Corner Metric*: A corner is a special two dimensional feature point which has high curvature in the region boundary. It can be detected by finding the local maximum in the corner metric. Corner metric is a feature that describes the possibility of the corner points in the different parts of the image.

The text appearing in the videos and images has greater edge strength and edge density compared to background of the image and hence has greater possibilities of having a corner points. This property of the text allows it to be detected by obtaining the corner metric on the video frames or the images.

There are many ways to obtain the corner metric in the image; we have used the corner detection method proposed by Shi and Thomasi [33], [35] (also referred as Kanade-Tomasi corner detection) since it is found to be the most accurate in detecting corners. Here we briefly explain the calculation of corner metric. For more technical details, see [33]. For a given image  $I(x, y)$ , the basic form of corner metric is shown in equation below.

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u + x, v + y) - I(u, v))^2$$

Here  $w(u, v)$  is a window function, which can be written in the matrix form as:

$$S(x, y) \approx (x, y)A(x, y)^T$$

Here  $A$  is given by:

$$A = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

Here  $I_x$  and  $I_y$  represents the partial derivative of  $I$ , the angle brackets denote averaging. A corner (or in general an interest point) is characterized by a large variation of  $S$  in all directions of the vector by  $(x, y)$  analysing the eigenvalues of  $A$ , this characterization can be expressed in the following way:  $A$  should have two large eigenvalues for an interest

point. Let  $\lambda_1$  and  $\lambda_2$  be the two eigenvalues of  $A$ , corner detection method computes a corner in the image if the following condition is satisfied:

$$\min(\lambda_1, \lambda_2) > \lambda$$

The above states that a corner is detected if both  $\lambda_1$  and  $\lambda_2$  has value that is greater than a predefined threshold value  $\lambda$ .

The sample input image is shown in figure 3 and the corresponding corner metric of the input image is shown in figure 4.

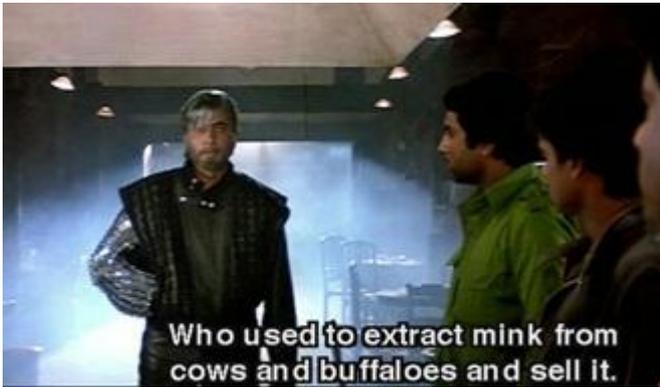


Fig. 3 Sample input image.



Fig.4 Corner Metric of the sample input image.

The resultant corner metric of the input image will be a binary image; the white pixels corresponding to the text region of the input image are the detected text pixels, while the remaining white pixels in the background are considered to be the noise pixels which must be removed for efficient text localization.

1) *Laplacian Filtering*: The Laplacian filter [34] is an isotropic filter i.e. the response is independent of the direction of the discontinuity in the image, such filters are rotation invariant.

Another important feature of the text appearing in the images and videos is the greater intensity discontinuity between the text and the image background, based on this property, the Laplacian filtering technique is used to detect the text as it highlights intensity discontinuities in an image and deemphasizes regions with slowly varying intensity levels. The Laplacian filter uses second order derivative (2-D) implementation for image sharpening, the mathematical equations of the Laplacian filter is given below:

$$\nabla^2 = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

The above equation can be implemented by using a Laplacian filter shown in the figure 5.

0	1	0
1	-4	1
0	1	0

Fig.5 Laplacian filter mask

The noise produced by the Laplacian method can be reduced by binarizing the filtered image using a global thresholding method. The figure 6 shows the result of the Laplacian filtering on the sample input image shown in figure 3.



Fig. 6 Filtered image of the sample input image.

The figure 6 shows the detected text pixels as well as the noise pixels which are removed in the next step.

2) *Multiplication*: Both the corner metric method as well as the Laplacian filtering method detects the artificial text but also produces noise in the background. For the text localization it is necessary to efficiently reduce the noise pixels.

Since the both the corner metric and Laplacian filtering make use of different properties of the text for detection, the noise produced by these two techniques will also be different from each other, based on this fact we combine (intersection) the result of corner metric and Laplacian filtering to form a single resultant binary image such that it contains only the common text pixels. Combining the resultant binary images of both the corner metric and the Laplacian filtering technique is done by multiplication. The figure 7 shows the multiplication result of the corner metric and Laplacian filtering results.

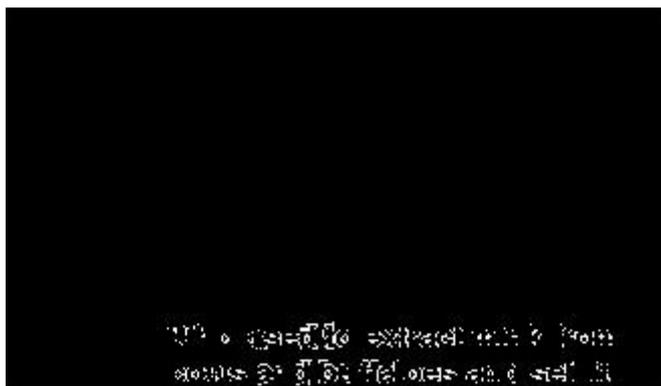


Fig.7 Multiplication result of corner metric and Laplacian filtering.

Apart from the noise reduction, the multiplication process also determines whether the given image or video frame does have an artificial text embedded in it. Multiplication process produces a blank image with zero white pixels if the data contains no text pixels. The shape of the detected text region is still irregular, and it needs to be refined into a aligned rectangle, this process is carried out in the text localization step.

### B. Text Localization

The binary image that is obtained from the text detection process is fed as an input to the text localization process, but the localization of the text is not done on binary image instead the done for the original input data by making use of the information present in the binary image obtained from the detection process.

In text localization process the binary is scanned for white (1) pixels in both X and Y axis, the first location of the first white pixel obtained in X axis is considered as  $X_{min}$  and the location of the last white pixel obtained in the X axis is considered as  $X_{max}$ , the same procedure is carried out in Y axis and the location of the first and the last white pixel in the Y axis is considered as  $Y_{min}$  and  $Y_{max}$  respectively. The pixels  $X_{min}$ ,  $X_{max}$ ,  $Y_{min}$  and  $Y_{max}$  are called as border pixels.

The border pixels are mapped onto the original input image /video frame and extended by two pixel position away from the text region; this forms a rectangular text region in the original image which is then cropped /segmented from the rest of the background. The result of the text localization process is shown in figure 8.



Fig.8 Localized text.

### C. Text Binarization

The final process of the text extraction mechanism is text binarization; the localized text obtained from the previous process is fed to text binarization process, the result of the binarization process should be in such a way that the text should be black (0) with white background (1), this condition is necessary for standard recognition algorithm to recognize the text from the binary image [32].

Since the text pixels should be black against the white background, the binarization technique like global thresholding works only when the colour of the text appearing in the video is known well in advance, it does not work globally for all the colours ranging from black to white. Hence determination of the text colour is necessary for efficient text binarization.

The process of determining the text colour is called seed selection, we assume that the colour of the text is unique in a given image or video frame, so determination of the single seed pixel provides sufficient information for binarization.

We have considered the that  $X_{min}$  obtained in the multiplication step provides the location of the beginning of the text, to make sure that the seed location falls inside the body of the text we added one position to  $X_{min}$  to get seed pixel location to get  $X_{seed}$ , hence  $X_{seed}$  can be obtained from  $X_{min}$  using the equation  $X_{seed} = X_{min} + 1$ . Then we map the  $X_{seed}$  to the original input image and obtain the intensity or grey scale value at that location, this obtained pixel value is considered to the seed pixel  $SP$ .

We consider  $SP$  to be pixel value (colour) of the text, but sometimes due to lossy compression of the image/video the pixel value of the entire text will not remain the same, there will be a slight deviation in the intensity value throughout the text, hence we cannot just rely on a single seed pixel, therefore we have to consider a range of seed pixels to obtain proper and effective binarization.

Let us consider the seed pixel values ranges from  $SP_{min}$  to  $SP_{max}$  which are minimum intensity seed value and maximum intensity seed value respectively. Along with  $SP_{min}$  and  $SP_{max}$ , all the pixels value which fall in-between  $SP_{min}$  and  $SP_{max}$  are considered to be the candidate seed values. The  $SP_{min}$  and  $SP_{max}$  can be obtained by subtracting and adding  $T$  to  $SP$ . The mathematical equation is given below:

$$SP_{min} = SP - T$$

$$SP_{max} = SP + T$$

Based on the minimum and maximum seed values we can perform the binarization, the binarization is performed on the localized text obtained from the previous step. We make use of double thresholding technique for binarization using  $SP_{min}$  and  $SP_{max}$  as the two threshold values. The pixels with intensity values less than  $SP_{min}$  and the pixels with intensity values greater than  $SP_{max}$  is converted to one (1) and the pixel values that fall in-between  $SP_{min}$  and  $SP_{max}$  is converted into zero (0). This results in a binary image with the black text pixels against the white background irrespective of the original colour of the text appeared in the input data. The figure 9 shows the result of text binarization performed on the localized text shown in figure 8.

Who used to extract milk from  
cows and buffaloes and sell it.

Fig. 9 Binarized text.

In our experiments we have the value of  $T$  as 5.

#### IV. EXPERIMENTAL RESULTS

The proposed solution has been tested on a number of videos and images of having artificial text in several languages, for different image resolutions. The figure 10 shows the average number of noise pixel distribution in all the three processes of text detection step calculated for 200 video frames which has been randomly selected from different videos.

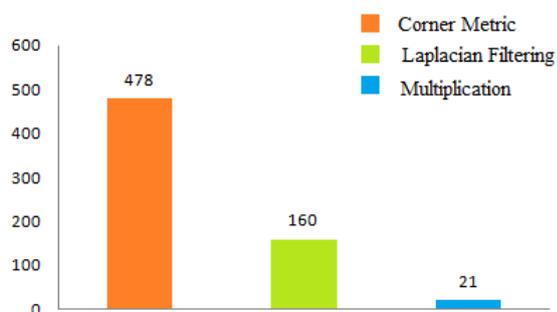


Figure. 10 Average noise pixel distribution in text detection step.

The figure 10 shows the effectiveness of the noise reduction step by multiplying the results of corner metric and Laplacian filtering, on an average the corner metric produces more noise pixels compared to the Laplacian filtering, as per the results the multiplication step removes the noise present in corner metric by 95.6% and Laplacian filtering by 86.9%.

The accuracy of the text detection and text localization of the proposed solution has been found to be 94.7%, and the accuracy of the text binarization step of the proposed solution has been found to be 91.2% for our test data.

#### V. CONCLUSION

The proposed solution effectively extracts the artificial multilingual text at different locations of different color and font. The experimental results have demonstrated the effectiveness and accuracy of the proposed solution.

#### REFERENCES

- [1] Marios Anthimopoulos, Basilis Gatos, Ioannis Pratikakis, "A two-stage scheme for text detection in video images", Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", 153 10 Athens, Greece.
- [2] Eliza Yingzi Du and Chein-I Chang, "Thresholding Video Images for Text Detection", 1051-4651/02 \$17.00 (c) 2002 IEEE.
- [3] A.K. Jain, and B. Yu, "Automatic text location in images and video frames", Pattern Recognition, 31(12): 2055-2076, 1998.
- [4] M.R Lyu and J.-Q. Song, "A comprehensive method for multilingual video text detection, localization, and extraction," IEEE Trans. Circuits and System for Video Technology, vol. 15, no. 2, pp. 243–255, 2005.
- [5] H. Li, D. Doermann, and O. Kia, "Automatic Text Detection and Tracking in digital Video", IEEE Trans. Image Processing, 9(1): 147-156, 2000.
- [6] Wu, R. Manmatha, and E.M. Riseman, "Finding Text In Images", in Proc. Of the 2nd Intl. Conf. on Digital Libraries. Philadelphia. PA. pp.1-10, July 1997
- [7] Li Sun, Guizhong Liu, Xueming Qian, Daping Guo, "A NOVEL TEXT DETECTION AND LOCALIZATION METHOD BASED ON CORNER RESPONSE", Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on June 28 2009-July 3 2009.
- [8] Ullas Gargi, David Crandall, Sameer Antani, Ryan Keener, Rangachar Kasturi, "A System for Automatic Text Detection in Video", Document Analysis and Recognition, 1999. ICDAR '99. Proceedings of the Fifth International Conference on: 20-22 Sep 1999
- [9] Haojin Yang, Bernhard Quehl, Harald Sack, "TEXT DETECTION IN VIDEO IMAGES USING ADAPTIVE EDGE DETECTION AND STROKE WIDTH VERIFICATION", Systems Signals and Image Processing (IWSSIP), 2012 19th International Conference.
- [10] Guangyi Miao, Qingming Huang, Shuqiang Jiang, Wen Gao: "Coarse-to-fine video text detection". ICME 2008: 569-572. 2007.
- [11] Palaiahnakote Shivakumara, Weihua Huang and Chew Lim Tan "Efficient Video Text Detection using Edge Features", Pattern recognition 2008, ICPR 2008.
- [12] Min Cai, Jiqiang Song, and Michael R. Lyu "A NEW APPROACH FOR VIDEO TEXT DETECTION", image processing 2002, ISSN: 1522-4880.
- [13] Marios Anthimopoulos, Basilis Gatos, Ioannis Pratikakis "A Hybrid System for Text Detection in Video Frames", Document Analysis System 2008.
- [14] Xi Jie, Xian-Sheng Hua, Xiang-Rong Chen, 2001. Liu Wenyin, HongJiang Zhang, "A Video Text Detection and Recognition System", IEEE International 22-25 Aug. 2001.
- [15] Hongming Zhang Wen Gao Xilin Chen Debin Zhao Neural Networks, "Learning informative features for spatial histogram-based object detection" 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on, 1806- 1811 vol. 3(2005).
- [16] V.Vijayakumar, "A Novel Method for Super Imposed Text Extraction in a Sports Video", International Journal of Computer Applications (0975 – 8887). Volume 15– No.1, February 2011.
- [17] Palaiahnakote Shivakumara, Weihua Huang, Trung Quy Phan and Chew Lim Tan, "Accurate Video Text Detection through Classification of Low and High Contrast Images", Pattern Recognition 2010.
- [18] P. Shivakumara, Trung Quy Phan and Chew Lim Tan, "New Fourier-Statistical Features in RGB Space for Video Text Detection", IEEE Transactions on CSVT, 2010. pp 1520-1532.
- [19] Shim, C. Dorai, and R. Bolle, Automatic Text Extraction from Video for Content-based Annotation and Retrieval, Proc. of International Conference on Pattern Recognition, Vol. 1, 1998, pp. 618-620.
- [20] E.K. Wong and M. Chen, "A new robust algorithm for video text extraction", Pattern Recognition, 2003, pp. 1397-1406.
- [21] Q. Ye, Q. Huang, W. Gao and D. Zhao, "Fast and robust text detection in images and video frames", Image and Vision Computing, 2005, pp. 565-576.
- [22] R. Lienhart and A. Wernicke, "Localizing and Segmenting Text in Images and Videos", IEEE Transactions on CSVT, 2002, pp 256-268.
- [23] T. Sato, T. Kanade, E. K. Hughes, and M. A. Smith, "Video OCR for digital news archive," in Proc. IEEE Workshop Content-Based Access Image Video Database, 1998, pp. 52–60.
- [24] M. Cai, J. Song, and M. R. Lyu, "A new approach for video text detection," in Proc. Int. Conf. Image Process., Rochester, NY, Sep. 2002, pp. 117–120.
- [25] C.G. Harris and M.J. Stephens, "A combined corner and edge detector," in Proceeding of the 4th Alvey Vision. Conference, 1988, pp. 147–152.
- [26] N. Otsu, "A threshold selection method from gray level histogram," IEEE Transactions on System, Man, Cybernetics, vol. 19, no. 1, pp. 62–66, 1978.
- [27] W. Niblack, An Introduction to Digital Image Processing, Englewood Cliffs, New Jersey: Prentice-Hall, 1986.
- [28] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," Pattern Recognition, vol. 33, no. 2, pp. 225–236, January 2000.
- [29] C.Wolf, J.-M. Jolion, and F. Chassaing, "Text Localization, Enhancement and Binarization in Multimedia Documents," in Proc. of the Int. Conf. on Pattern Recognition, 2002, vol. 2, pp. 1037–1040.
- [30] C. M. Thillou and B. Gosselin, "Color text extraction with selective metric-based clustering," Computer Vision and Image Understanding, vol. 107, pp. 1–2, July 2007.
- [31] Z. Zhou, L. Li, and C. L. Tan, "Edge based binarization for video text images," in Proc. of 20th Int. Conf. On Pattern Recognition, Singapore, 2010, pp. 133–136.
- [32] Haojin Yang, Bernhard Quehl, "A SKELETON BASED BINARIZATION APPROACH FOR VIDEO TEXT RECOGNITION", 13th international workshop on image analysis for multimedia interactive services 23-25 may 2012, Dublin Ireland.
- [33] Jiambo Shi, Carlo Tomasi, "Good Features to Track", IEEE conference on Vision and Pattern Recognition (CVPR94)Seattle, June 1994.
- [34] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Third Edition, published by Pearson Education, Inc, publishing as Prentice Hall, copyright © 2008.
- [35] B. D. Lucas and T. Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision", International Joint Conference on Arti\_cial Intelligence, pp. 674-679, 1981