

# Extraction and Labeling Images from PDF Documents

M.S. Kavya and B.G. Geetha

**Abstract---** The main motivation for extracting the object from pdf file is accessibility problem. Current world is enclosed with enormous masses of digital visual information. The rapid development in technologies and images urging for development of robust and efficient object detection technique. The focal point of the most reported literatures is object detection and its applications. Single object from an image can be easily identified. An object can also be identified from a cluster of images simultaneously. The paper discusses about the various techniques for object recognition and a method for identifying object from a multiple image.

---

## I. INTRODUCTION AND SCOPE

The rapid increase in the diversity and availability of electronic information led to further processing requirements, in order to enhance and accessing the images. This problem is even more admissible for audio visual information where data has to be searched, queued and processed. The objects need to be searched are stored in a database. Initially the pdf file is read by using the code. Then the object need to be identified is read as an input. This particular input is compared with the pdf image. The objects in the pdf images are extracted by using the specific techniques. These extracted images are compared with the input image. It checks the pixels of both the images. Size component is not at all important factor in this technique. If the object is found in the particular image from the cluster of image that are extracted from the pdf it will label and display the particular object.

---

*M.S. Kavya, PG Scholar, Department of CSE, K.S.Rangasamy College of Technology, Tiruchengode. E-mail:kavyasakthive1253@gmail.com*

*B.G. Geetha, Professor, Head of the Department, Department of CSE, K.S.Rangasamy College of Technology, Tiruchengode.*

## II. LITERATURE SURVEY

**Cosmin Ancuti, Philippe Bekaert, Codruta O. Ancuti, Christophe De Vleeschouwer;** describes the enhancement of the images captured underwater and degraded due to scattering and absorption. To evade that the sharp weight map conversions create artefacts in the low frequency components of the reconstructed image, we also adapt a multiscale fusion strategy.

The strategy builds on the integration principle and does not require further additional information than the single original image. The approach is able to enhance an enormous range of underwater images (e.g. different cameras, depths, light conditions) with peak accuracy, being able to get back important faded features and edges. Moreover, for the first time, it demonstrates the utility and relevance of the proposed image enhancement technique for several challenging underwater computer vision applications.

**Qingsong Zhu, Ling Shao, Jiaming Mai;** describes the misty removal is a challenging problem due to its unwell-posed nature. By creating a linear model for designing the scene depth of the hazy image under this novel earlier and learning the parameters of the model with a supervised learning method, the depth information can be well recovered.

A novel linear color attenuation earlier, based on the difference between the brightness and the saturation of the pixels within the misty image. By creating a linear model for the scene depth of the murky image with this simple but muscular prior and learning the parameters of the model using a supervised learning method, the depth information can be well recovered. By means of the depth map the scene radiance of the hazy image can be recovered easily.

**Ming-Wei Wu, Wujie Zhou, Lu Yu, Weiwei Qiu, Yang Zhou;** describes a powerfully unsighted quality evaluation technique for SCIs and NSIs based on a dictionary of learned localized and globalized quality attributes. Initially, a localized glossary is developed using locally normalized image patches and conventional K-means clustering. By using the local glossary, the learned local quality attributes can be obtained using a locality-constrained linear coding with max pooling. In order to extract the learned global quality attributes, the histogram representations of 0s and 1s patterns are joined to form a global dictionary. The collective representation algorithm is used to efficiently code the learned global quality features of the misshape images using this glossary. At last, kernel-based support vector regression is used to mix these attributes into an overall quality score. Substantial experiments involving the proposed rating techniques that are verified in comparison with most relevant metrics, the blind metric yields importantly higher stability in line with subjective fidelity ratings.

A constructive unsighted quality predictor for misshape images by incorporating the learning of both localized and globalized quality attributes. The innovation of the research occupy in merging the opposite behaviours of the locality-constrained linear coding based local quality features and joined representation based global quality attribute to attain a fused representation for images.

**Mitra Fatemi, Martin Vetterli, Arash Amini, Loic Baboulaz;** describes visual signals captured as digital images. Thus, the reconstructed image is conditioned to regenerate the same samples as well as forming a shape (bilevel) image. It initially formulates the reconstruction technique by minimizing the shape perimeter over the set of consistent binary shapes. The reducibility that guarantees equality between the two problems. The instance that the reducibility property effectively sets a requirement on the minimum sampling density. Examine the performance of the various numerical methods.

It also describes about the problem of remodelling a continuous-domain shape image from the samples in a gray-scale discrete image. It is necessary to the introduction of pixels in a way that generates a binary image.

**Gabriela Ghimpe, teanu, Marcelo Bertalmío, Thomas Batard, Stacey Levine;** describes removing noise from an image is a basic task for correcting defects generated during the asset process of a real world scene and its generation on a display is due to physical and technological limitations. This is also useful as a pre-processing stage in order to increase the results of top level applications.

It enables any noise removing method to take more into account the local geometry of the image to be denoised by protecting the floating frame describing the graph of a scaled version of the image. It is easier to improve the performance of three noise removing methods of different types: a patch-based method, a local variational method and a method combining a patch based approach with a filtering in spectral domain approach, display the consistency of our methodology. The elements have different geometric meaning, one shall awe whether they should not rather be denoised with different denoising methods.

**Kyong Hwan Jin, Michael T. McCann;** describes a deep convolutional network for reciprocal problems with a focal point on biomedical imaging. The structure of the model is based on U-net, with the addition of residual learning. The particular approach was activated by the loop structure of several biomedical inverse problems. The constraints on a linear operator that shield that its normal operator is a convolution.

The proposed method demonstrated compelling results on synthetic and real data. It compared favourably to state-of-the-art iterative reconstruction on the two more realistic datasets. The next training, the total time of the proposed network per one image is under a second.

**Jianbing Shen, Wenguan Wang, Ling Shao;** describes the deep video saliency model has two modules, namely static saliency network and dynamic saliency

network, which are modeled for capturing spatial and temporal statistics of dynamic scenes. The saliency estimates from the static saliency network is incorporated in the dynamic saliency network.

It enables our method to automatically learn the way of fusing static saliency into dynamic saliency detection and directly produce final spatiotemporal saliency results with less computation load. Furthermore, we proposed a novel data augmentation technique for synthesizing video data from still images, which enables our deep saliency model to learn generic spatial and temporal saliency and prevents overfitting. Additionally, our model is very efficient with a frame rate of 2fps on a GPU.

**Amin Zheng, Dinei Florencio, Gene Cheung;** describes the advent of depth sensing practical application means that the extraction of object contours in images-a common and important pre-processing step for later higher-level computer vision tasks like object detection and human action recognition-has become easier.

A separate two-stage approach that first denoises then encodes contours lossily. Adopting a joint approach, we propose a burst error model that models typical errors encountered in an observed string of directional edges. We then formulate a rate-constrained maximum a posteriori (MAP) problem that trades off the posterior probability of an estimated string given with its code rate. We design a dynamic programming (DP) algorithm that solves the posed problem optimally, and propose a compact context representation called total suffix tree(TST)that can reduce complexity of the algorithm dramatically. To the best of our knowledge, we are the first in the literature that study the problem of joint denoising /compression of image contours and offer a computation-efficient optimization algorithm. Experimental results show that our joint denoising/compression scheme can reduce bitrate by up to 18% compared to a competing separate scheme at comparable visual quality.

## REFERENCES

- [1] A. Zheng, G. Cheung and D. Florencio, "Joint Denoising/Compression of Image Contours via Shape Prior and Context Tree", *IEEE Transactions on Image Processing*, 2018.
- [2] M. Fatemi, A. Amini, L. Baboulaz and M. Vetterli, "Shapes from pixels", *IEEE transactions on image processing*, Vol. 25, No.3, Pp.1193-1206, 2016.
- [3] C.O. Ancuti, C. Ancuti, C. De Vleeschouwer and P. Bekaert, "Color balance and fusion for underwater image enhancement", *IEEE Transactions on Image Processing*, Vol. 27, No. 1, Pp.379-393, 2018.
- [4] G. Ghimpețeanu, T. Batard, M. Bertalmío and S. Levine, "A decomposition framework for image denoising algorithms", *IEEE transactions on image processing*, Vol. 25, No. 1, Pp.388-399, 2016.
- [5] Q. Zhu, J. Mai and L. Shao, "A fast single image haze removal algorithm using color attenuation prior", *IEEE Transactions on Image Processing*, Vol. 24, No. 11, Pp.3522-3533, 2015.
- [6] W. Wang, J. Shen and L. Shao, "Video salient object detection via fully convolutional networks", *IEEE Transactions on Image Processing*, Vol. 27, No. 1, Pp.38-49, 2018.
- [7] Z. Li, J. Zheng, Z. Zhu, W. Yao and S. Wu, "Weighted guided image filtering", *IEEE Transactions on Image processing*, Vol. 24, No. 1, Pp.120-129, 2015.
- [8] K.H. Jin, M.T. McCann, E. Froustey and M. Unser, "Deep convolutional neural network for inverse problems in imaging", *IEEE Transactions on Image Processing*, Vol. 26, No.9, Pp.4509-4522, 2017.
- [9] W. Zhou, L. Yu, Y. Zhou, W. Qiu, M.W. Wu and T. Luo, "Local and Global Feature Learning for Blind Quality Evaluation of Screen Content and Natural Scene Images", *IEEE Transactions on Image Processing*, Vol. 27, No.5, Pp.2086-2095, 2018.
- [10] S.D. Khirade, and A.B. Patil, "Plant disease detection using image processing", *International Conference on Computing Communication Control and Automation*, Pp. 768-771, 2015.
- [11] S.K. Chachra, Z. Xue, S. Antani, D. Demner-Fushman and G.R. Thoma, "Extraction and labeling high-resolution images from pdf documents", *Document Recognition and Retrieval XXI (Vol. 9021, p. 90210Q)*, International Society for Optics and Photonics, 2014.
- [12] W. Du, Y. Wang and Y. Qiao, "Recurrent Spatial-Temporal Attention Network for Action Recognition in Videos", *IEEE Transactions on Image Processing*, Vol. 27, No. 3, Pp.1347-1360, 2018.