Study on Soft Edge Transition Correction using Multilayer MRC Model

Umesh P. Akare
Research Scholar
Priyadarshini College of Engg.,Nagpur

N.G.Bawane, Ph.D
Principal
S.B.Jain College of Engg. & Management ,Nagpur

ABSTRACT
The mixed raster content (MRC) image represents a compound image which is a superposition of layers. This image model is very efficient for representing sharp text and graphics onto a background. In this binary mask layer is used. The problem occurs when one deals with scanned data and soft edges. These edges are neither shown as a background nor as a foreground. To detect segmented soft edges and a method to correct and sharpen the image within the MRC model is highly required. This paper presents details on this aspect while discussing preprocessing of the compound images with the MRC standard model.

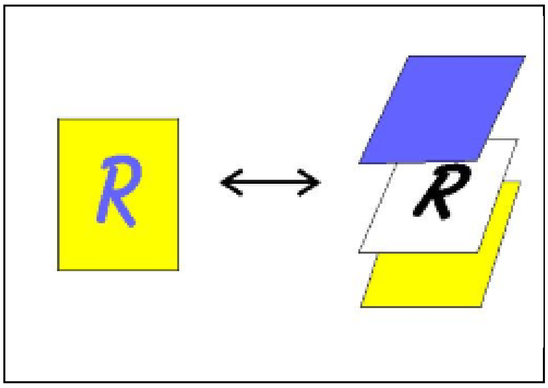
Keywords
Mixed Raster Content (MRC) Model, Preprocessing, Segmentation, Edge Transition

1. INTRODUCTION
In recent years, compound document image compression has become an important research field due to economical reasons for compressing huge amount of scanned documents. A compound or mixed document contains text, graphics and pictures together. They are a very common form of documents. The examples where one can find these are magazines, brochures, web-sites etc. Due to the nature of these two image classes (text or graphics vs. pictures), their compression strategies logically involves multiple compression schemes and a region segmentation i.e. classification method [1, 2]. It is challenging task to compress such type of documents with the traditional compression scheme, which are generally suitable with a particular image type and application specific. Different resolutions are needed for textual and natural images [3].Textual an image requires higher resolution in space color values of pixels. It requires less tonal resolution but sufficient spatial resolution to preserve the features of the text [4, 5]. Document images usually have important visual details like irregularity in font, color and texture of the paper. These are important particularly in historical documents and handwritten text [6].

1.1 Compound Documents
There are basically two forms of electronic documents i.e. vectorial and raster. It is easy to compress vectorized document since its each object can be compressed individually but to compress rasterized documents is having real challenges. Compression algorithms are generally developed to suit particular image type, its characteristics and application for which it is implemented. No single algorithm is found best suitable for all types of images and applications.
It is important to preserve the edges as well as shapes of the characters to facilitate the reading especially when compressing text and line art. When it comes to compound documents, in order to cope with the differences between text and continuous tone images, different compression algorithms may be applied to each of the regions of the document image. To achieve the goal, some segmentation strategy has to invariably be used to decide which regions are to be encoded under which strategy. Another important parameter of a document compression system for compound documents is its imaging model. One can separate the image into different regions of interest and compress each region accordingly. In this case, the imaging model follows space segmentation where each decompressed region can be imaged into the document concurrently. Also, one can generate multiple image layers, compress each one separately and then image all the planes into one.

1.2 Mixed Raster Content (MRC)
Compound raster documents have typically been compressed as a single image. Different compression schemes may be applied to each of the regions of the document. This is a way like multiple-coder based algorithm generally works. The mixed raster content (MRC) imaging model is a multi-layer and multiresolution representation of a compound document. The first step of MRC compression is the layer segmentation algorithm. The 3-layer MRC model represents an image as two image layers, first representing Foreground or FG and second representing Background or BG and a binary Mask image layer. The Mask layer is helpful to reconstruct the final image from the FG/BG layers. This model can be properly understood with the help of figure1.

Fig 1: 3Layer decomposition of compound document image



The foreground plane is to be poured through the mask plane onto the background plane. The original image is to be decomposed into layers; each layer can then be processed separately and subsequently can be compressed using different algorithms. Layers may have different dimensions with offsets associated with them [7]. A binary mask layer, foreground layer, and background layer are differently shown below in the MRC standard model (fig.2).

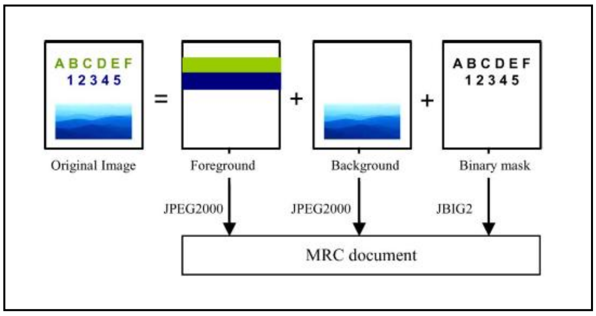


Fig 2 : MRC Standard Model

The binary mask represents the assignment of each pixel to the foreground layer by a "1" represented by black or the background layer by a "0" represented by white color, respectively. Generally text regions are classified as foreground and picture regions are classified as background. Then each layer is encoded individually using an appropriate encoder [8]. MRC has been proposed and accepted for several standards as well as been used in several products.

1.3 Compression Standard

DjVu, LuraDocument and DigiPaper have been already implemented in products as MRC-based compression approaches [6, 9, 10]. DjVu is one of the existing state–of-art compression and segmentation method. It uses wavelet base encoder to encode the foreground and background layers and uses k-means based clustering method to generate the mask layer. Many modern approaches have compared their performance with that of DjVu [11, 12].Its compression performance is better than those single encoder compression methods and other mentioned products. It causes color diffusion from adjacent color regions having uniform color distribution. Text segmentation is the most critical step in MRC encoding. DjVu have problem relating to text segmentation. The quality of decoded image and reducing the bit rate are primary concern. Segmentation classifies textual and graphical regions to create binary mask layer. The misclassification reduces the text quality and readability of the document text. In Farsi and Arabic document images where such textual details frequently occur and the misclassification have found important effect on text quality, readability. Jaggy artifacts have found more undesirable effects. These artifacts make the word segmentation difficult. Word segmentation is preferably used in almost all Farsi and Arabic OCR systems [13].These artifacts also can affect up to omitting thin parts of the text and hence create difficulties in recognition of text and also distort the image quality. This leads to erroneous OCR results.The multi-encoder MRC-based method produces better image quality. It has not any ringing effect and has well preserved the edges quality compared to JPEG2000. Although as a result of losing the anti-aliasing effect, the DjVu decoded image exhibits jaggy artifacts around the text edges. The same are found in other ones also. This problem is not desirable in high-quality MRC document coding .Many MRC-based methods can yield much higher compression ratios than conventional color image compression methods. But the binary representation used tends to distort fine document details, such as thin lines and text edges [7]. Edge reconstruction is the prominent problem of most MRC based compression methods including the DjVu. The MRC model could achieve high compression ratios with keeping high reconstructed image quality is reported in [2, 3, 14]. ITU standard (T.44) originally proposed for color facsimile. It has also been proposed as the JPEG2000 architecture framework for the compound image file format. A key approach in the T.44 implementation of MRC is that all the three (or more) layers is to be compressed independently by a suitable encoder. This approach [2, 3] has advantages in its simplicity and modularity although it increases coding overhead. The binary mask can only represent discontinuous transitions between text, line art, and background colors. This sets an intrinsic limitation on the standard MRC model. This type of hard transition can introduce substantial artifacts. The real documents often contain continuous tone transitions between regions. These continuous tone transitions are helpful in anti aliasing which allow encoding at lower dpi. This can further reduce bit rate. Zaghetto et al. has tried to get rid of this problem. He experimented with blurring boundary regions using a Gaussian filter [15]. But this lowers the quality of the reconstructed image. One can try to solve this problem by modeling the edge regions by a polynomial function of a low order. We review state-of-the-art technologies on the subject while focusing our attention on the mixed raster content (MRC) multi-layer approach.

2. LITERATURE REPORTED

2.1 Main Approach

There are two main approaches through which we can decompose compound images. They are based on object segmentation or on region classification [16]. The object decomposition identifies text and graphics objects by pouring text or graphics ink through the mask plane onto the BG plane. Here the mask image layer would contain text characters, line art and filled region and the foreground layer contains the colors of text letters and graphics. In region classification approach, regions containing text and graphics are identified. Then it is represented in a separate (foreground) plane. The whole region is represented in the foreground plane which includes the spaces in between letters also. Here the mask formed is very uniform with large patches indicating the text and graphics regions. The BG plane contains the document background, complex graphics and/or continuous tone pictures. The text itself is contained within the FG plane. If compression ratio is the main motivation, then it is often more useful to employ object segmentation.□□Region□□ classification avoids redundancy and is more natural for efficient compressors to deal with segmented image data compression. However, it requires modifying MRC, on the□□ other hand, employs stock compression only. Document segmentation is a specific binarization procedure which keeps only textual objects. Improper threshold selection can causes blotches, streaks, erasures, merges, fractures, deformations in the character shapes [17]. The resulting quality and compression ratio of a MRC document encoder is mainly dependent on the segmentation algorithm by which the binary mask is computed [18].

2.2 Document Segmentation Algorithms

Most existing compression algorithms for document images can be roughly classified as block-based approaches and layer-based approaches. Block-based approaches segment non-overlapping blocks of pixels into different classes. Then it compresses each class differently according to its characteristics. In layer-based approaches, it partitions a document image into different layers, like background and the foreground layer. Then, each layer is coded as an image independent from other layers. Most layer-based approaches are based on three-layer (foreground/mask/background) representation proposed in the ITU's Recommendations T.44 for mixed raster content (MRC). The foreground layer contains the color of text and line graphics while the background layer contains pictures and background.Both the quality and bitrate of an MRC document depends upon segmentation method selected. For example, if a text component is not properly detected the text edges will be blurred by the background layer encoder. On the other hand, if non text is erroneously detected as text will also cause distortion. It may be in the form of false edge artifacts and the excessive smoothing of regions assigned to the foreground layer. An erroneously detected text can also increase the bit rate required for symbol-based compression methods such as JBIG2. This occurs due to erroneously detected and unstructured non text symbols which are not efficiently represented by JBIG2 symbol dictionaries. The performance of a document compression system is directly influenced by segmentation algorithm. A suitable and a good segmentation approach is that which can lower the bit rate as well as lower the distortion too. Most damaging artifacts are often caused by misclassifications. Some segmentation algorithms those proposed for document compression use features extracted from the discrete cosine transform (DCT) coefficients to separate text from picture.

Several segmentation methods have been proposed in the past on layer-based compression approach. It is viewed as the procedure combining document segmentation through thresholding (or binarization) and modification procedures. Most existing segmentation methods have directly or indirectly implemented this scheme. Out of many methods available for image thresholding only a specific group [17, 18, 19] methods are found suitable for document images. In [17, 19] the existing thresholding methods are evaluated and then concluded that for document applications, techniques from locally adaptive and shape categories are found the top thresholding methods.Conventional natural image compression standards such as JPEG or JPEG2000 are usually used for encoding the foreground and the background layer. The mask layer is a bi-level textual image. This layer mostly encoded using pattern matching based textual compression methods. Overall compression performance and the bi-level mask layer compression are critically affected by document segmentation. The pattern matching (PM) scheme was originally presented in [20].As a standard, MRC is intended to employ standard compressors as well. JPEG is the most widely adopted and used compression method for color images. It is well suited to compress both the FG and BG layers. Because of the large number of variables involved, we want to keep the MRC model as simple as possible .it does not include resizing the layers before compression. JPEG 2000 is a newer compression standard designed to upgrade the original JPEG .It is based on wavelet transforms and contextual coding of bit planes, achieving a near State-of-the-art performance by the time it was devised[21].Final pixels are dependent on various layers and therefore they can be unused while processing the layer individually. In the segmented image, pixels assigned to the BG will be marked as "don't care" on the FG and vice-versa. Preprocessor can be used to replace such pixels by any color in order to enhance the compression. The problem of data filling over the redundant (unused) data to enhance compression has been used in [7].The compression algorithm and resolution used for a given layer would be matched to the layer's content, testing for improved compression. It reduces distortion in visibility. In [16], author has presented threshold based layer segmentation algorithm and iterative data filling algorithm for redundant regions. In [8], video compression standard H.264/AVC is applied as a still image compression tool. It outperforms previous state-of-art coders such as JPEG2000. In [22], a segmented layer image mask separation algorithm (SLIm) for compression of document images has been designed. Unlike DjVu and TIFF-FX, it produces symmetric separation between foreground and background. This proves better particularly for comic books. The main advantages of the mask separation in SLIm are its algorithm simplicity, dependency on a few parameters, its symmetry and speed. It has promising refinements as it uses linear approximation of the foreground and background instead of constants when computing the mask. In [23], it proposes a general quadrilateral-based framework for image segmentation. Quadrilaterals are first constructed from an edge map. Neighboring quadrilaterals with similar features of interest are then merged in them together to form regions. This enables the elimination of local variations and unnecessary details for merging. Each segmented region is accurately and completely described by a set of quadrilaterals. The proposed algorithm shows that the regions obtained are segmented into multiple levels. It shows that the proposed algorithm performs better than three other segmentation techniques such as seeded region growing, constrained gravitational clustering and K-means clustering. It offers an efficient description of the segmented objects for content-based applications. A robust multiscale segmentation algorithm for both detecting and segmenting text in complex documents is presented in [24].The approach found suitable for those documents containing noisy backgrounds, background gradations, reversed contrast text, and varying text size. Authors approach differs in that it integrates a stochastic model of text structure and context into a multiscale framework so that to best meet the requirements of MRC document compression. This method minimizes false detections of unstructured non text components. This creates artifacts and ultimately increases bit-rate. But this method accurately segments true text components of varying size and with variation in backgrounds. This algorithm can achieve higher decoded image quality at a lower bit-rate than generic algorithms for document segmentation. The author uses two algorithms that are applied in sequence: the cost optimized segmentation (COS) algorithm and the connected component classification (CCC) algorithm. Block wise segmentation COS algorithm works on cost optimization. Gray level or color document images are converted to binary image by the COS algorithm. The resulting binary image typically contains many false text detections. The CCC algorithm further improves the accuracy of the segmentation through detecting non text components (i.e., false text detections) in a Bayesian framework which incorporates a Markov random field (MRF) model of the component labels.

3. EDGE TRANSITION AND ITS PREPROCESING

In the scanned image, the edges of text and graphics are not as sharp and we represents them as "soft edges." MRC uses a binary mask and therefore is suitable to represent text and graphics with sharp edges. Since, the mask plane is binary and the edge transitions are gradual, it is not possible to contain all the background in one plane and the entire foreground in another. The mask determines what is BG or FG. Note that the transition is included in both planes. If, in the BG plane, we replace the parts corresponding to FG with the average of the BG plane for increased compression, we would not end up with a flat line, but with a spike. Same thing occurs with the FG plane[25].

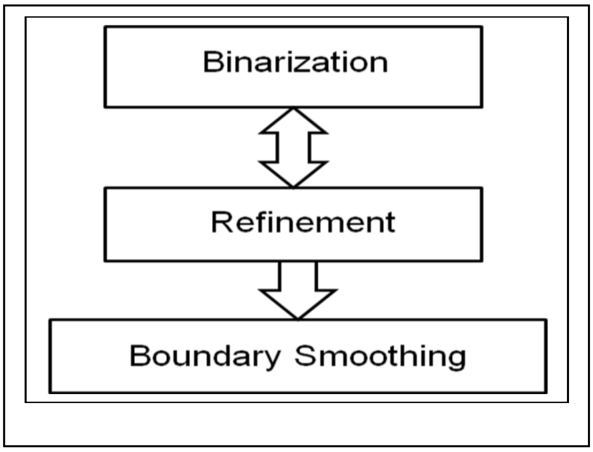


Fig 3: Segmentation Flow Chart

Fig.3, shows three processes sequentially such as binarization, refinement, and boundary smoothing, can become part of the segmentation approach. The binarization block binarizes the input document image based on the transition regions in the input image. The resulting binary image may have some non-textual parts. These parts are then removed by the refinement process. It uses some prior knowledge on properties of textual objects [26].

.

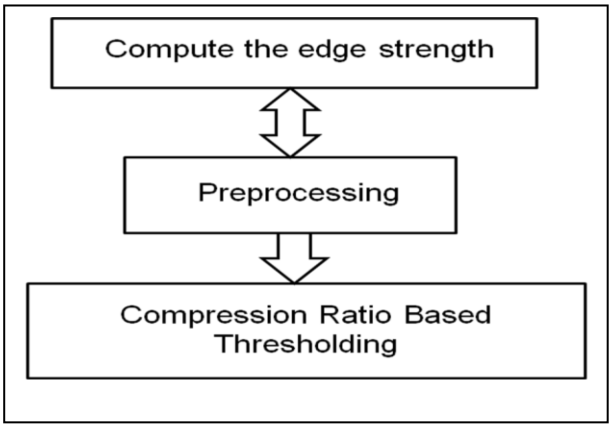


Fig 4: Binarization Procedure

The binarization procedure is shown in the stages of flowchart of Fig.4. It finds the image transition regions between textual and background regions. Edge transition regions have relatively high edge strength [26].

An improved data filling algorithm for the redundant regions is presented in [27]. The methods have developed to counter balance the effects of soft edges in MRC compression of scanned images. A method to sharpen only those soft edges that might cause the halo effect was developed. At the decoder, the reconstructed image edges are re-softened. The re-softening can be done by sending a halo location map to the decoder or by estimating the softness of the original image and recreating it at the decoder. Edge sharpening map and iteratively parameterize the original edge at the encoder is implemented through the algorithm at the encoder. This is then used to reconstruct the original soft edges at the decoder. This can yield gains in PSNR within compression range of interest. In [1], the algorithm is also developed which iteratively estimates the original edge softness through Gaussian filter. It generates halo location map. The filter parameters in conjunction with map are used to reconstruct soft edges at the decoder. It shows comparative improvements in performance for low bit rates. We have to somehow estimate the transition of the image edges .The images are sharpened to accommodate the mask and to reconstruct soft edges [1, 27].In another approach the author uses the same basic idea but describes three new features that improve the compression performance. The real bitrate is used to determine the pre/post-processing parameters. The Mask layer and the edge sharpening/softening map used with improved quality. At last layer resolution change has been explored. It shows that the method provides higher quality as well as it can also yield up gains in PSNR as compared to single coder approaches, in the compression ranges of interest [21]. In [7], a method called resolution-enhanced rendering (RER) is proposed. It is combinely optimizes the MRC encoder and decoder to achieve low distortion rendering of edge transitions in the MRC binary mask layer. The method works on the basis of adaptively dithering the mask layer of a three-layer MRC encoding which produces the intermediate tone levels required for low distortion rendering.

4. CONCLUSION

In this paper we discussed predominant techniques for the compression of mixed raster compound documents in addition to that we have reviewed the methods those developed to counter balance the effects of soft edges in MRC compression of scanned images. When the image is scanned, the edges of text and graphics are not as sharp as expected. We refer it as soft edges. As the selector plane is binary and the edge transitions are gradual, it is not possible to contain all the foreground in one plane and all the background in another. The mask decides what is BG or FG. The transitions are included in both planes. This effect creates a halo around the text. This damages quality of compression. Since it occurs inside we could not depends only on data filling techniques. The halo can be more pronounced depending on how the mask was built. These halos cannot be removed with data filling. Therefore we have to forcefully change the data itself. We cannot assume that all image edges can cause the problem. As well as we cannot assume that all masks transitions are causing halo in scanned material. This effect occurs only when transition of the masks coincides just with the edges of the image. This proves that a sharp mask transition is used to model a soft edge transition, which leads the halo. Hence one has to find first where the mask transitions coincide with the image edges. In this paper, we have discussed many segmentation schemes applicable mainly to MRC document images. Refining is expected to improve the quality of the images. Lastly we have discussed the edge transition problem and suggest various ways of preprocessing schemes to overcome it.

5 REFERENCES

[1] de Queiroz, R.L., Fan, Z., Tran, T.D., 2000, Optimizing block-thresholding segmentation for multilayer compression of compound images. IEEE Trans. Image Proc. 9(9), 1461–1471

[2]Ricardo L. de Queiroz,Oct.1999, Compression of Compound Documents. Proc. IEEE international Conference on Image Processing, ICIP, Kobe, Japan, 25PS1.1.

[3] Feng, G., Bouman, C.A., 2006, High-quality mrc document coding. IEEE Trans. Image Proc. 15(10), 3152–3169

[4] de Queiroz, R., Buckley, R., Xu, M., 2000, Mixed raster content (MRC) model for compound image compression. In: Proceedings images. IEEE Trans. Image Proc. 9(9), 1461–1471.

[5] Lam, E.Y.,2004, Compound document compression with model-based biased reconstruction. J. Electron. Imaging 13(1), 191–194

[6]Bottou, L., Haffner, P., Howard, P.G., Simard, P., Bengio, Y., LeCun, Y.,1998, High quality document image compression with DjVu.J. Elec. Imaging 7(3), 410–425

[7] Ricardo L. de Queiroz, 2006, Pre-Processing for MRC Layers of Scanned Images. ICIP ,pp-3093-3096

[8]Eri Haneda, Charles A. Bouman, June 2011, Text Segmentation for MRC Document Compression. IEEE Transactions on Image Processing, Vol. 20, No. 6.

[9] Barthel, K. U., Partlin, S. M., Thierschmann, M., 2000, New technology for raster document image compression. In: Part of the IS&T/SPIE Conference on Document Recognition and Retrieval VII, San Jose, CA, vol. 3967, pp. 286–290

[10]Thierschmann, M., Barthel, K. U., Martin, U. -E., 2001, A scalable DSP architecture for high-speed color document compression. In: Kantor, P.B., Lopresti, D.P., Zhou, J. (eds.) Document Recognition and Retrieval VIII. Proceedings of SPIE, vol. 4307, pp. 158–166

[11]Wu, B.-F., Chiu, C.-C., Chen, Y.-L.,2004, Algorithms for compressing compound document images with large text/background overlap. IEE Proc. Vis. Image Signal Process. 151(6), 453–459

[12]Haneda, E., Yi, J., Bouman, C. A., 2007, Segmentation for MRC compression. In: Eschbach, R., Marcu G.G. (eds.) Color Imaging XII: Processing, Hardcopy, and Applications. Proceedings of SPIEIS&T Electronic Imaging, vol. 6493, pp. 252–262

[13]Sarfraz, M., Zidouri, A., Nawaz, S. N.,2005, On Offline arabic character recognition. In: Sarfraz M. (ed.) Computer-Aided Intelligent Recodgnition: Techniques and Applications. Wiley, NY

[14]Fan, Z., Jacobs, T.,2005, Segmentation for mixed raster contents with multiple extracted constant color areas. Proc. SPIE-IS&T Electron. Imaging 5667, 251–262

[15] Zaghetto, A., de Queiroz, R. L.,2008, Iterative pre and post processing for MRC layers of scanned documents. In: Proceedings of IEEE International Conference on Image Processing, ICIP, San Diego, CA, USA, Oct.

[16] R. L. de Queiroz, D. Mukherjee, 2008, MRC Compression of Compound Documents Using Threshold Segmentation, Iterative Data-filling and H.264/AVC-INTRA. Sixth Indian Conference on Computer Vision, Graphics & Image Processing, ICVGIP.

[17]Sezgin, M., Sankur, B.,2004, Survey over image thresholding techniques and quantitative performance evaluation. J. Elec. Imaging 13(1), 146–165

[18]Haneda, E., Yi, J., Bouman, C. A.,2007, Segmentation for MRC compression. In: Eschbach, R., Marcu G.G. (eds.) Color Imaging XII: Processing, Hardcopy, and Applications. Proceedings of SPIEIS&T Electronic Imaging, vol. 6493, pp. 252–262

[19]Trier, O.D., Taxt, T.,1995, Evaluation of binarization methods for document images. IEEE Trans. Pattern Anal. Mach. Intell. PAMI 17(3), 312–315

[20] Ascher, R.N., Nagy, G., 1974, A means for achieving a high degree of compaction on scan-digitized printed text. IEEE Trans. Comput. 23, 1174–1179

[21] R. L. de Queiroz, 2005 Compressing Compound Documents, in The Document and Image Compression Handbook, edited by M.Barni,Marcel-Dekkar,

[22]Patrice Y. Simard, Henrique S. Malvar, James Rinker, and Erin Renshaw,2004, A Foreground/Background Separation Algorithm for Image Compression, Proceedings of the Data Compression Conference (DCC'04)

[23]Ronald H.Y. Chung, Nelson H.C. Yung, Paul Y.S. Cheung, 2005, An Efficient Parameterless Quadrilateral-Based Image Segmentation Method, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 27, No. 9,pp-1446-1458

[24]Alexandre Zaghetto, Ricardo L. de Queiroz, 2008, Iterative Pre- and Post-Processing for MRC Layers of Scanned Documents. ICIP.pp-1009-1012

[25]Barthel, K. U., Partlin, S. M., Thierschmann, M., 2000, New technology for raster document image compression. In: Part of the IS&T/SPIE Conference on Document Recognition and Retrieval VII, San Jose, CA, vol. 3967, pp. 286–290

[26] Hadi Grailu · Mojtaba Lotfizad · Hadi Sadoghi-Yazdi, 2009, Farsi and Arabic document images lossy compression based on the mixed raster content model , IJDAR 12:227–248