21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

A Novel Scheme of Orientation and Scale mapped RDC (OS-RDC) to improve compression in Document Images ensuring quality preservation

Apurba Das, Remya R
Imaging Tech Lab., HCL Technologies Ltd., Chennai, India
apurba.das@hcl.com, remya_r5@yahoo.co.in

Abstract

Repeated appearance of any block of spatial data in document images can be cached and encoded single time to get good compression ratio. This Reusable Document Component (RDC) can replicate the blocks of each redundant image at the receiver side at different positions but with same size and orientation. We have proposed a novel algorithm of Orientation Scale mapped RDC (OS-RDC) which can identify and cache repeated image blocks even if they are of different size and orientations. Both the inter-page and intra-page redundancies are addressed ensuring significant quality preservation.

Keywords: *Reusable Document Component (RDC); caching; compression; Fuzzy C-means; orientation and scale map.*

1. Introduction

In communication [11], printing [12] and digital library based applications [13], betterment in the process of document image encoding and compression for efficient utilization of bandwidth, time and storage respectively are always state-of-art research areas [1], [3], [4], [6], [15]. Repeated appearance of any block of spatial data in document images can be cached and encoded a single time to get a good compression ratio. This Reusable Document Component (RDC) can replicate the blocks of the image at the receiver side at different positions but with same size and orientation. We have proposed a novel algorithm of Orientation Scale Mapped RDC (OS-RDC) which can identify and cache repeated image blocks even if they are of different size and orientation. Both the inter-page and intra-page redundancies are handled ensuring significant quality preservation. O. E. Kia et. al. have used RDC for text blocks [14]. In our approach we have segmented the image blocks from the entire document and marked by orientation and scale map to each peer match. To ensure quality in document image, we have proposed a new Fuzzy C- means based algorithm of adaptive document image compression where we can tune the compression ratio and quality in either way by the membership values of each block (of size 8x8) belonging to lossy and lossless regions of compression. In the printing industry it is always prescribed to use lossless compression for texts [7], [8], [9], [10].

In the next section, we have described our proposed scheme of OS-RDC. The method of finding the relative orientation and scale map from peer matched repeating image blocks is described. In section 3, we have presented some experimental results where the OS-RDC is applied over different well-known compression algorithms to achieve better compression ratios in each case. In section 4, we have proposed one adaptive compression algorithm which will ensure the quality preservation of document images. We have shown that, the integration of OS-RDC with state-of-art compression algorithms ensures quality preservation, significantly. Finally in section 5 we have concluded our findings with some directions for future work.

2. Orientation and Scale Mapped RDC

The researchers in the document image domain have always prescribed to avoid texts in lossy encoded or compressed manner as even a small distortion may change the appearance of a text significantly by missing corners or edges. A Fuzzy C-means (FCM) clustering method is used to obtain the image and non-image region from the document. The clustering is based on the features such as color variance, edge strength and high frequency to low frequency ratio. Therefore we have concentrated on the image blocks to form Reusable Document Components (RDC). In our proposed scheme we shall cache the largest of the repeating images with different scale and orientation and replace their position only by relative orientation and scale (OS) map, which is discussed in the later sub-sections. By adding the OS map, we can extend the concept of caching of recurring blocks to OS-RDC to address the problem more robustly. During reconstruction, the images are obtained using the scale and orientation map. The pseudo-code of the proposed scheme is as follows.

```
Iseg← image segments
NIseg←nonimage segments
numIseg←total number of image segments
t← threshold
Segment document image into Iseg and NIseg using FCM
Mark all image segments as 'not matched' segments
Consider all Iseg
for  imsegi =  1 to numIseg
      for imsegj = (imsegi+1) to numIseg
            if image segments imsegi and imsegj are
            'not matched' segments
                  Match using scale & rotation invariant algorithm
                  if (matching score >= t)
                        map← Find the scale and orientation map
            for each pair imsegi and imsegj
                  Using map, find the largest image segment and
            cache it
                  Mark matched smaller segment as 'matched'
                  Store OS map and pointer to the largest image
            in place of smaller image
                        end
               end
      end
end
```

2.1 Formation of OS-RDC by Matching

To find the peer matched between already identified image blocks, we have used a psycho-visually inspired rotation and scale invariant matching algorithm [2], which is designed keeping the concept of DoG pyramids in mind like SIFT [5]. There are four major stages of computation to generate the set of image features as:

1. **Scale-space extrema detection:** A difference-of-Gaussian (DoG) function is applied to identify potential interest points that are invariant to scale and orientation. A possible scale-space kernel is Gaussian function. Therefore, the scale space of an image is defined as a function, $L(x,y,\sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x,y,\sigma)$ with an input image $I(x,y)$:

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \quad (1)$$

where * signifies convolution and

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

To efficiently detect stable keypoint locations in scale space, Lowe [5] suggested obtaining scale-space extrema in the Difference-of-Gaussian image $D(x,y,\sigma)$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k:

$$\begin{aligned}D(x,y,\sigma)&=(G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y)\\&=L(x,y,k\sigma)-L(x,y,\sigma)\end{aligned} \quad (3)$$

2. **Keypoint localization:** At each candidate location, a detailed model is fitted in order to determine the location and scale. The keypoints are selected based on measures of their stability.

3. **Orientation assignment:** One or more orientations are assigned to each keypoint location based on local image gradient directions.

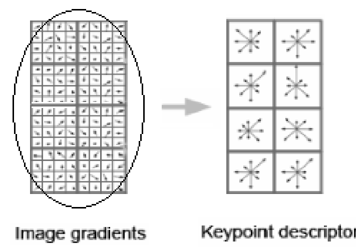


Figure 1. Formation of feature descriptor

4. **Keypoint descriptor**: The local image gradients are measured at the selected scale in the region around each keypoint. The descriptor is formed from a vector containing the values of all the orientation histogram entries. From the computed gradient of each key-point the 2×4 neighborhood (Figure 1), we have considered 8 orientation bins. Hence feature descriptor of length 2×4×8=64 is formed for each key-point [2]. In classical SIFT [5], Lowe used a 4x4 neighborhood, but Apurba et. al has shown [2], with respect to the response of visual cortical column, elliptic orientation measurement is sufficient. As the feature vector size is reduced, the computational cost is also less. As seen in the 1st column of Table 1, the system itself can identify and detect the proper position of the peer match of interest in presence of scale and rotation in terms of matching the descriptors of the points. If the accumulated matching score between each pair of feature points is higher than the predefined threshold (T=0.5), we declare them a "matched pair".

2.2 Orientation and Scale Map

A novel technique for quantifying the relative occupancy of a target object is presented in this section using the following pseudo-code.

```
Iseg← image segment
NIseg←nonimage segments
numIseg← total number of image segments
t←threshold
k←keypoints
```

tk← terminal keypoints (TKP)
d1← distance between tk in left image used in matching
d2← distance between tk in right image used in matching
simsegij← scale map of imsegi and imsegj
θimsegij← orientation map of imsegi and imsegj

```
for imseg = 1  to numIseg
    k← Find the keypoints for each image segment imseg
    Find the direction of variance of keypoint distribution.
    θimseg← Orientation normalization: If the direction is not
aligned to the vertical axis, rotate each  Iseg anticlockwise
aligned to the vertical axis
end
Mark all image segments as ‘unmatched’ segments
for  imsegi = 1 to numIseg
   for  imsegj = i+1 to numIseg
        Compare each pair of unmatched image segments
        if (matching score >= t)
            kl1 , kl2← Find TKP of left image
            kr1 , kr2← Find TKP of right image
            d1← distance between kl1 and kl2
            d2← distance between kr1 and kr2
            simsegij← 1/(d2/d1)
            θimsegij← abs(θimsegi – θimsegj)
            if(d1>=d2)
                 Left image segment is larger
            else
                 Right image segment is larger
            end
            Mark “matched” smaller image segments
        end
    end
end
```

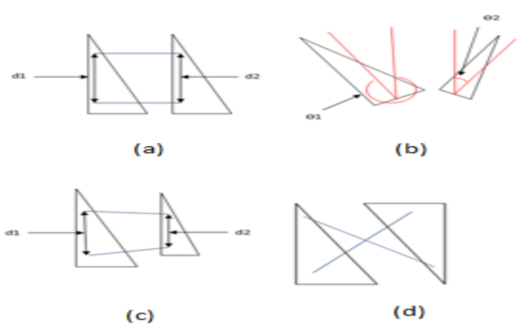
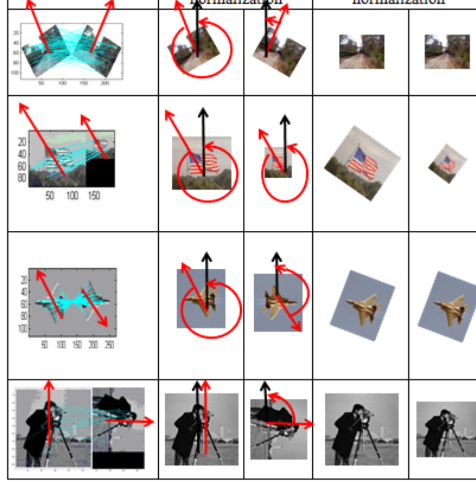


Figure 2. Peer match between images of (a) same size and orientation, (b) different size and orientation, (c) different size, (d) different orientation

For peer match between same size and orientation images, the terminal matched lines are parallel, and then the angle θ is zero degrees. If the angle is zero degrees and ratio d1:d2 is 1:1, the two images are exactly of same size and orientation (Figure 2(a)). The orientation map (O-map) is the difference between θ1 and θ2 (Figure 2(b)). Then the orientation normalization is performed by rotating both images such that both of their principle axes (axis joining terminal match points) align to the vertical axis. The relative angle difference forms the orientation map. Then the scale map(S-map) is calculated by finding d1:d2 (Figure 2(c).) from orientation normalized image pair. Some experimental results are given in Table 1.

Table 1: Orientation and Scale map



Matching	Before orientation normalization		After orientation normalization		O-map	S-map
					300	1:1
					0	1:4
					180	1:1
					-90	1:2

3. Improvement in compression by OS-RDC

We have tested and integrated our algorithm with multiple well-known compression techniques and tested on different document image databases formed by a random set of pages [16].

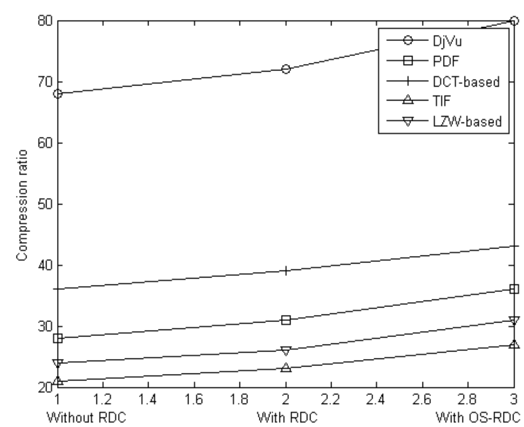


Figure 3. Comparison of compression ratio of various techniques

In classical RDC, while encoding the image in compressed format, the same image with different scale and rotation are cached separately, whereas in our proposed scheme, we have cached only the largest image and represented the scale and rotation varying images by reference and OS-map. The results are depicted in the graph shown in Figure 3.

4. Quality preservation in OS-RDC

The proposed scheme does not further degrade the quality of the encoded image with respect to the classical compression. The measures of PSNR are depicted in Table 2, which supports the claim, ensuring un-erroneous reconstruction of the image blocks from the reference and OS map. Along with the other algorithms, we have also tested OS-RDC integrated with the proposed FCM based Quality Preserving Compression (QPC) algorithm and compared the PSNR degradation by OS-RDC in Table 2.

Table 2. PSNR values of images of various techniques

	Without RDC	With OS-RDC
DjVu	22.742	22.700
PDF	10.828	10.666
DCT-based	10.711	10.589
TIF	10.644	10.533
LZW-based	10.563	10.466
QPC	44.100	44.003

5. Conclusion

In the present paper, a novel method of orientation and scale mapped RDC (OS-RDC) is presented. By integrating the proposed scheme of caching with a set of image compression, we could achieve our improved compression ratio without compromising the quality. As the RDC is scale and orientation invariant and tunable in terms of fuzzy memberships, the concept can be applied to different domain of image based communication, representation and encryption.

References

[1] Behl P., Chaudhury S. and Lall B., "Image Analogy Based Document Image Compression", *Third National Conference onComputer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, pp. 223-226, December 2011, Hubli, Karnataka

[2] A. Das, A. Roy and K. Ghosh, "Proposing a CNN Based Architecture of Mid-level Vision for Feeding the WHERE and WHAT Pathways in the Brain", *SEMCCO, Proceedings, Part I. Lecture Notes in Computer Science, Springer*, Vol. 7076, pp. 559-568, December 2011, Visakhapatnam, Andhra Pradesh

[3] A. Aung, B. P. Ng, and C. T. Shwe, "A New Transform For Document Image Compression*," in Proc. 2009 7th IEEE International Conference on Information, Communications and Signal Processing*, pp. 1-5, December 2009, Macau

[4] L. Bottou, P. Haffner, P. Howard, P. Simard, Y. Bengio, and Y. LeCun, "High Quality Document Image Compression with DjVu", *Journal of Electronic Imaging*, Vol.7, No. 3, pp. 410-425, July 1998

[5] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, November 2004

[6] R. Kountchev, V. Todorov, M. Milanova and R. Kountcheva, "Documents Image Compression with IDP and Adaptive RLE", *32nd Annual Conference on IEEE Industrial Electronics, IECON*, pp. 2361 – 2366, November 2006, Paris

[7] R. Radescu, "Transform methods used in lossless compression of text files, " *Romanian Journal of Information Science and Technology(ROMJIST)*, Vol.12, No. 1, pp. 101-115, 2009, Bucharest

[8] R. Radescu, "Lossless Text Compression Using the LIPT Transform", *Proceedings of the 7th International Conference Communications 2008 (COMM2008)*, pp. 59-62, June 2008, Bucharest, Romania

[9] F. Awan and A. Mukherjee, "LIPT: A lossless Text Transform to Improve Compression", *Proceedings of International Conference on Information and Theory: Coding and computing, IEEE Computer Society*, April 2001, Las Vegas, Nevada

[10] R. Franceschini, H. Kurse, N. Zhang, R. Iqbal and A. Mukherjee, "Lossless, Reversible Transformations that Improve Text Compression Ratios", IEEE Transactions on Multimedia Systems, June 2000

[11] Müldner T., Leighton G. and Diamond J., "Using XML compression for WWW communication", *Proceedings of the IADIS WWW/Internet 2005 Conference*, 2005

[12] M. G. Ramos and R. L. de Queiroz, "Classified JPEG Coding of Mixed Document Images for Printing", *IEEE transactions on image processing*, Vol. 9, No. 4, April 2000

[13] A. Moffat and I. H. Witten, "A Compression-Based Digital Library", *DESIDOC Bulletin of Information Technolog*y, Vol. 17, pp. 31-41, 1998

[14] O. E. Kia and D. S. Doermann, "Residual Coding in Document Image Compression", *IEEE Transactions on Image Processing*, Vol. 9, No. 6, pp. 961-969, June 2000

[15] N. C. Francisco, N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria and V. M. M. Silva, "Scanned Compound Document Encoding Using Multiscale Recurrent Patterns", *IEEE Transactions on Image Processing*, Vol. 19, No. 10, pp. 2712-2724, October 2010

[16] Gonzalez and Woods, "Digital Image processing" , Second Edition , Prentice Hall, Upper Saddle River, New Jersey 07458, 2002

[17] A. Das, S. Mitra and S. Parua, "Hierarchical Clustering on Texture Statistics for Search Space Reduction of Large Scale Face Recognition", *IEEE International Conference on Image Information Processing (ICIIP)*, November 2011