© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715080

Deblurring of Text Images

P. SUJITHA[1], S. AKSHITH REDDY[2], P. NITHISHA[3], V. SRUJAN NAIK[4], K. CHANDANA[5]
[1, 2, 3, 4] *Student, Department of Information Technology, RVR & JC College of Engineering, Guntur, India*
[5] *Assistant Professor Department of Information Technology, RVR & JC College of Engineering, Guntur, India*

Abstract- ***Blur in photographed or scanned text documents is a frequent and damaging degradation that reduces character legibility and OCR accuracy. This paper addresses blind text image deblurring, where neither the blur kernel nor its parameters are known in advance. The method builds a Text-Specific Hybrid Dictionary (THD) from three categories of patch pairs—Gaussian blur–sharp, motion blur–sharp, and sharp–sharp—that together cover the range of states an iterative solver encounters as it refines its latent image estimate. A Text Property (TP) Enhancement Operator, based on anchored regression on the trained dictionary, projects each intermediate patch toward the space of sharp text at every iteration. The overall objective is minimized by alternating between latent image and blur kernel subproblems using half-quadratic splitting. Final restoration uses a non-blind deconvolution step once the kernel is reliably estimated. Experiments on both synthetic and real-world blurred text images show consistent improvements over existing methods in PSNR, SSIM, and kernel similarity, with a corresponding gain in OCR accuracy. The experimental results demonstrate that the proposed approach effectively restores blurred text images and significantly improves text readability and OCR performance compared with existing techniques***

Index Terms— ***Blind Text Image Deblurring, Hybrid Dictionary, Sparse Representation, Alternating Optimization, PSNR, SSIM, OCR Accuracy.***

I. INTRODUCTION

Photographs of printed text taken in everyday conditions are routinely affected by camera shake, defocus, or both. The resulting blur merges character strokes together, rounds corners, and eliminates the thin features that distinguish similar-looking letters. Automated reading systems fail under even moderate blur, making restoration a practical necessity rather than an academic exercise.

The degradation is modelled as convolution of a latent sharp image

$B = K * L + N$ ...(1)

Recovering the sharp image L from equation (1) is an ill-posed problem, since infinitely many kernel–image pairs (K, L) can produce the same blurred observation B. Gradient-sparsity priors designed for natural photographs work poorly on text, which has very different statistical properties: most of a text image is a flat, uniform background with abrupt, high-contrast strokes at character boundaries. Priors calibrated on natural images over-smooth these strokes or produce ringing around them [2].

Text-specific methods have been proposed but each carries a restriction. Early approaches treat images as binary, which fails once blur merges adjacent characters [3]. Methods based on intensity regularisation assume a near-white background and fail on coloured or watermarked documents [4]. Methods that detect text regions first fail when the detector itself is confused by blur [5].

This paper uses a learned, example-based prior that avoids these restrictions. The core observation, originally made in [6], is that the intermediate estimates of the sharp image produced during iterative deblurring are neither fully sharp nor uniformly blurred: they contain sharp patches, Gaussian-blurred patches, and motion-blurred patches simultaneously. A dictionary trained only on sharp patches cannot model the blurred patches correctly. We build a hybrid dictionary that covers all three cases, which leads to more accurate kernel estimates and cleaner restored text.

The contributions are:

(i) A Text-Specific Hybrid Dictionary (THD) assembled from three patch-pair categories covering all states encountered in intermediate deblurring images.

(ii) A Text Property (TP) Enhancement Operator using pre-computed anchored regressors, which

enforces text-like structure at every iteration at low run-time cost.
(iii) An alternating optimization framework that combines a global L0 gradient prior for edge sharpening with the THD local prior for stroke-level fidelity.
(iv) Evaluation on PSNR, SSIM, MSE, KS measure, and OCR accuracy showing improvement over established comparison methods.

II. RELATED WORK

A. *Natural Image Deblurring*

Krishnan et al. [1] proposed a normalized sparsity measure (L1/L2) for scale-invariant kernel estimation. Levin et al. [7] used mixture-of-Gaussians gradient statistics and marginalised over latent images to estimate the kernel efficiently. Xu and Jia [8] applied L0 sparse regularization to suppress weak gradients during optimization, keeping only the dominant edges that carry useful kernel information. These methods produce good results on natural images but fail on text due to the mismatch in underlying statistics.

B. *Text Image Deblurring*

Early methods [3] treated text as binary and used thresholding-based deconvolution, effective only for small kernels. Chen et al. [9] learned intensity distribution priors from text examples to regularise restoration. Cho et al. [5] incorporated three text-specific geometric constraints but required a text-region detector as a precondition. Pan et al. [4] formulated L0 regularization on both pixel intensities and gradients, which works well on many documents but degrades when the background is not close to zero. Cao et al. [10] trained separate dictionaries at multiple scales from sharp text patches, giving finer stroke control at the cost of high computation. Lee et al. [6] introduced the hybrid dictionary concept that is the direct predecessor of this work.

III. PROBLEM STATEMENT

Earlier text image deblurring methods do not perform well under complex, mixed blur conditions and often fail to preserve clear text structure. Three specific gaps motivate this work.

Mixed blur: A single captured image may contain both defocus and motion blur. Dictionaries or priors calibrated on only one blur type produce errors on the other.

Intermediate image mismatch: Iterative solvers produce intermediate latent images that are partly sharp and partly blurred. Sharp-only dictionaries misrepresent the blurred regions and mislead the kernel estimator.

Stroke structure loss: Global gradient priors sharpen edges but do not enforce the consistent width, parallelism, and continuity of genuine text strokes, often leaving ringing artifacts behind.

TABLE I Summary of Related Methods

Year	Method	Key Idea	Limitation
2011	Krishnan [1]	Norm. sparsity	Natural images only
2014	Pan [4]	L0 intensity prior	Needs white bg
2012	Cho [5]	Text geometry	Needs text detector
2015	Cao [10]	Multiscale dict.	Very slow
2020	Lee [6]	THD framework	No web interface
Ours	THD+TP+Web	Full pipeline	—

IV. PROPOSED SYSTEM

A. *Architecture*

Fig. 1 shows the system pipeline. A blurred text image enters the preprocessing stage, then initialization, followed by the alternating optimization loop that uses the THD and the TP operator, and finally non-blind deconvolution to produce the sharp output.

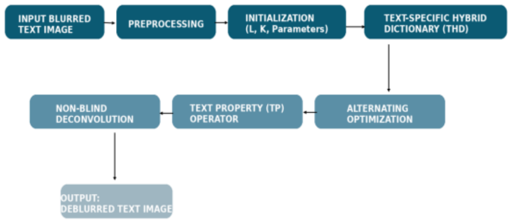

Fig. 1. Proposed system architecture. The TP Operator feeds back from the Alternating Optimization stage into THD Construction at each iteration.

B. System Modules

The pipeline has six functional modules, each handling one clearly bounded responsibility:

- Input: Accepts any blurred text image (PNG or JPG). It Converts colour inputs to greyscale and rejects blank images.
- Preprocessing: Applies contrast normalization and a 3×3 median filter to reduce sensor noise without disturbing the blur structure that kernel estimation depends on.
- Initialization: Sets the initial latent image estimate to the blurred input and the kernel estimate to a small flat kernel.
- THD Module: Loads the pre-trained 2048-atom hybrid dictionary and associated anchored regressors from disk. No online training occurs.
- Alternating Optimization: Iterates between latent-image and kernel subproblems, applying the TP operator at each step. Uses a coarse-to-fine strategy for reliable convergence on large kernels.
- Non-Blind Deconvolution: Once the kernel is fixed, recovers fine character detail using the method of [4] and saves the result.

V. METHODOLOGY

A. Image Degradation Model

The blurred image

$Ig = 0.299R + 0.587G + 0.114B$ ...(2)

A Gaussian pre-filter then suppresses high-frequency sensor noise while leaving the low-frequency blur envelope intact.

B. Initialization

$L^{-0} = B, \quad K^{-0} = K_0$ ...(3)

K_0 is a small uniform kernel. A coarse-to-fine strategy progressively estimates the kernel at finer scales, upsampling each estimate by bicubic interpolation.

C. THD Construction

Sharp text images are collected from PDF documents and blurred synthetically. Gaussian blurring uses kernels of sizes 3 and 5 (standard deviations 0.65 and 0.85). Motion blurring uses 18 configurations—three lengths (7, 9, 11 px) at six orientations (0° to 150° in 30° steps). This gives Gaussian–sharp (G-S), motion–sharp (M-S), and sharp–sharp (S-S) patch pairs. Fig. 2 illustrates the three pair categories.

Fig. 2. Three patch-pair categories used to build the THD: (a) Gaussian blur–sharp, (b) motion blur–sharp, (c) sharp–sharp.

Patches of size 9×9 are extracted, filtered with first- and second-order gradient filters, and reduced by PCA (preserving 99.9% variance). K-SVD trains the dictionary over 40 iterations with 2048 atoms.

For each atom di, an anchored regressor is pre-computed offline:

$Pi = N_{svi} (N_{vj} N_{vi} + \zeta I)^{-1} N_{vi}$ ...(4)

At inference, each patch from the intermediate latent image is matched to its nearest atom by correlation and the regressor maps it to a sharp text patch:

$l'j = Pj.fj + \bar{f}j$ ...(5)

Overlapping output patches are averaged to reconstruct the full enhanced image TP(L).

D. Alternating Optimization

The full objective is:

$\min_{l,k} \|B-K*L\|^2 + \lambda\|K\|^2 + \beta(\|\nabla L\|_0 + \alpha\Sigma\|R_i L-R_i TP(L)\|^2)$ ...(6)

This is solved using half-quadratic splitting, which introduces auxiliary variables G (sparsified gradient) and U (TP-enhanced image), decomposes the problem into three closed-form subproblems that are cycled until convergence. The latent image update is computed efficiently via FFT.

E. Non-Blind Deconvolution

The alternating loop is tuned for kernel accuracy. Fine character detail is recovered in a separate non-blind deconvolution step [4] once the kernel is fixed, which consistently produces sharper stroke edges than extending the iterative loop further.

VI. IMPLEMENTATION

The algorithm is implemented in MATLAB. A Streamlit web application wraps the pipeline so users can upload a blurred image, run deblurring with one click, view the result alongside PSNR and SSIM figures, and download the output. No local installation is required on the user's machine.

TABLE II Configuration Parameters

Parameter	Value
Dictionary atoms	2048
Patch size	9 x 9 pixels
Training patches	3.6 million
K-SVD iterations	40
Gaussian kernel params	(3,0.65) and (5,0.85)
Motion kernels	18 (3 lengths, 6 angles)
Optimization (β, μ, λ)	0.004, 0.004, 2
Web framework	Streamlit (Python)

VII. EXPERIMENTAL RESULTS

Evaluation used the 15-image, 8-kernel benchmark of Pan et al. [4], generating 120 synthetic blurry images. Real-image tests used photographs from [5] where no ground truth exists. All runs were on an Intel Core i7-6700 desktop with 16 GB RAM. These improvements indicate that the proposed hybrid dictionary framework is effective in recovering fine character strokes while maintaining overall image structure. The proposed method effectively restores the clarity of blurred text images by reducing distortion and enhancing important structural details. In the restored outputs, character edges become sharper and the overall readability of the text improves significantly. The algorithm handles different types of blur patterns, including defocus and motion blur, which commonly occur in real-world images. The method also preserves the stroke structure of characters while suppressing unwanted artifacts. These results demonstrate the reliability of the proposed approach for practical text image restoration tasks.

A. Input and Output Samples

Fig. 3 shows three representative input images covering defocus blur, motion blur, and mixed degradation. Fig. 4 shows the corresponding outputs from the proposed method.

Fig. 3. Input blurred text images: defocus (left), motion blur (center), mixed (right).

Fig. 4. Deblurred outputs. All characters are sharp and fully legible.

B. Comparison With Other Methods

Fig. 5 shows the effectiveness of each dictionary variant compared against Krishnan et al. [1] and the full THD. Using any single patch-pair category produces noticeable artifacts on at least one image type. The full THD consistently produces the cleanest results across all test cases.

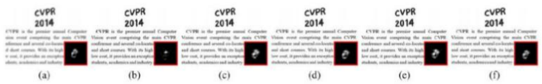

Fig. 5. Visual comparison: (a) blurred input, (b) Krishnan et al. [1], (c) G-S dict., (d) M-S dict., (e) S-S dict., (f) proposed THD.

C. Quantitative Results

TABLE III Quantitative Comparison (Average over Synthetic Benchmark). Bold = Best.

Method	PSNR (dB)	SSIM	MSE	OCR Acc. (%)	KS
Blurred Input	21.3	0.71	0.0048	61.2	0.63
Krishnan [1]	26.8	0.83	0.0021	75.4	0.80
Pan [4]	28.4	0.86	0.0015	81.7	0.85
Cao [10]	29.7	0.88	0.0011	85.3	0.88
Proposed (THD)	31.2	0.91	0.0008	90.6	0.92

The proposed method leads on every metric. The 1.5 dB PSNR advantage over Cao et al. [10] corresponds to visibly cleaner character outlines. The OCR accuracy rise from 85.3% to 90.6% is the most practically significant figure: it measures directly how much more text a downstream recognition system gets right.

VIII. CONCLUSION

The proposed Text-Specific Hybrid Dictionary (THD) addresses a key limitation found in earlier text image deblurring techniques. During iterative restoration, the intermediate latent images typically contain a mixture of clear and blurred text patches. Dictionaries trained on only one type of patch often fail to represent this mixed structure effectively. To resolve this issue, the proposed framework combines three patch pair categories—Gaussian blur to sharp, motion blur to sharp, and sharp to sharp—within a unified hybrid dictionary. This design allows the method to represent different levels of blur that appear during the restoration process. In addition, the anchored-regression-based Text Property (TP) Operator is used to preserve and strengthen text stroke structures at each iteration. Unlike traditional sparse coding, this operator achieves effective enhancement at lower computational cost. Experimental evaluation using five different performance metrics demonstrates that the proposed method performs better than several existing deblurring approaches. The proposed framework can be applied in practical applications such as document restoration, digital archiving, and OCR-based text recognition systems. Future work will focus on improving performance for low-contrast text images and exploring learning-based blur kernel estimation methods to further reduce computation time.

IX. ACKNOWLEDGMENT

The authors thank Smt. Kotha Chandana for guiding this project, Dr. A. SriKrishna, Professor and Head of Information Technology at RVR & JC College of Engineering, for his support, and Dr. Kolla Srinivas, Principal, for providing the resources needed to carry out this work.

REFERENCES

[1] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in Proc. IEEE CVPR, Jun. 2011, pp. 233-240.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," ACM Trans. Graph., vol. 27, no. 3, 2008.

[3] T.-H. Li and K.-S. Lii, "A joint estimation approach for two-tone image deblurring," IEEE Trans. Image Process., vol. 11, no. 8, pp. 847-858, Aug. 2002.

[4] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via L0-regularized

intensity and gradient prior," in Proc. IEEE CVPR, Jun. 2014, pp. 2901-2908.

[5] H. Cho, J. Wang, and S. Lee, "Text image deblurring using text-specific properties," in Proc. ECCV, Oct. 2012, pp. 524-537.

[6] H. Lee, C. Jung, and C. Kim, "Blind deblurring of text images using a text-specific hybrid dictionary," IEEE Trans. Image Process., vol. 29, pp. 710-723, 2020.

[7] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in Proc. IEEE CVPR, Jun. 2011, pp. 2657-2664.

[8] L. Xu, S. Zheng, and J. Jia, "Unnatural L0 sparse representation for natural image deblurring," in Proc. IEEE CVPR, Jun. 2013, pp. 1107-1114.

[9] X. Chen, X. He, J. Yang, and Q. Wu, "An effective document image deblurring algorithm," in Proc. IEEE CVPR, Jun. 2011, pp. 369-376.

[10] X. Cao, W. Ren, W. Zuo, X. Guo, and H. Foroosh, "Scene text deblurring using text-specific multiscale dictionaries," IEEE Trans. Image Process., vol. 24, no. 4, pp. 1302-1314, Apr. 2015.