

Contextual Restoration of Severely Degraded Document Images

Jyotirmoy Banerjee, Anoop M. Namboodiri, and C.V. Jawahar
Center for Visual Information Technology
IIIT, Hyderabad - 500032, India

email:{jyotirmoy@research.,anoop@,jawahar@}iiit.ac.in

Abstract

We propose an approach to restore severely degraded document images using a probabilistic context model. Unlike traditional approaches that use previously learned prior models to restore an image, we are able to learn the text model from the degraded document itself, making the approach independent of script, font, style, etc. We model the contextual relationship using an MRF. The ability to work with larger patch sizes allows us to deal with severe degradations including cuts, blobs, merges and vandalized documents. Our approach can also integrate document restoration and super-resolution into a single framework, thus directly generating high quality images from degraded documents. Experimental results show significant improvement in image quality on document images collected from various sources including magazines and books, and comprehensively demonstrate the robustness and adaptability of the approach. It works well with document collections such as books, even with severe degradations, and hence is ideally suited for repositories such as digital libraries.

1. Introduction

Degradations in document images result from poor quality of paper, the printing process, ink blot and fading, document aging, extraneous marks, noise from scanning, etc. The goal of document restoration is to remove some of these artifacts and recover an image that is close to what one would obtain under ideal printing and imaging conditions. The ability to restore a degraded document image to its ideal condition would be highly useful in a variety of fields such as document recognition, search and retrieval, historic document analysis, law enforcement, etc. The emergence of large collections of scanned books in digital libraries [1, 19] has introduced an imminent need for such restorations that will aid their recognition or ability to search.

Images with certain known noise models can be restored using traditional image restoration techniques such as Median filtering, Weiner filtering, etc. [9]. However, in prac-

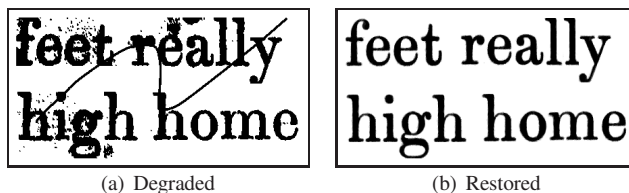


Figure 1. Portion a vandalized degraded document and the result of our restoration process.

tice, degradations arising from phenomena such as document aging or ink bleeding cannot be described using popular image noise models. Document processing algorithms improve upon the generic methods by incorporating document specific degradation models [20] and text specific content models [22, 16].

Approaches that deal with highly degraded documents (see figure 1) take a more focused approach by modeling specific types of degradations. For instance, ink-bleeding or backside reflection is one of the main reasons for degradation of historic handwritten documents. Huang *et al.* [12] combines the degradation model and the document model into a powerful MRF-based optimization framework [8, 13]. To achieve generic restoration of carbon copy documents, Cao and Govindaraju [3] used a document content model. The model consisted of a set of 5×5 binary patches, trained using high quality data, which is used for restoring noise and removal of rulings on the paper. Gupta *et al.* [10] used a patch based alphabet model to remove blurring artifacts for license plate images using a camera. The authors use an MRF based optimization to find the most likely noise free patch that generated the observations. All the above approaches consider specific instances of restoration of a single document image, and are solved by combining prior knowledge of documents with noisy observations.

In this paper, we approach document restoration in a different, and useful setting. We consider the problem of restoration of a degraded ‘collection of documents’ such as those from a single book. Such a collection of documents,

arising from the same source, is often highly homogeneous in the script, font and other typesetting parameters. The availability of such a uniform collection of documents for learning allows us to:

- Do robust learning of a tight model of the document content even in presence of severe degradations, as one can discard data that is potentially noisy.
- Do accurate parameter estimation from multiple evidences, as the amount of data available after discarding highly noisy parts is still considerable.
- Adapt to a large varieties of documents in various fonts, styles and scripts, as our model is exclusively learned from the input collection itself.

Given that we can learn an accurate and exact model of the documents content, we leverage it to compute the most probable estimate of the underlying content during document restoration. We frame the restoration process as a maximum a posteriori estimate computed from the learned document model prior and the noisy observed data in a Markov Random Field framework. Our formulation enables us to incorporate a larger context into the inferencing process, thus providing us with the ability to restore highly degraded documents.

The proposed approach is far more powerful than traditional approaches in restoring highly degraded documents as it relies on learning of a document model specific to the input. It can handle severe degradations including cuts, merges, ink blobs, or even vandalized documents. To achieve this, we address the problems of learning high quality priors and that of robust restoration in a flexible MRF-based optimization framework [7].

2. Restoration by Learning

The process of restoration proceeds in two stages. In the first stage, a set of ideal patches, x_i , that can occur in the restored document are estimated, along with the spatial relationship between them. This constitutes a probabilistic document model that is specific to the input document. The most likely set of patches that generated the observed patches, y_i , is estimated in the second stage, using an MRF framework. Figure 2 shows the construction of the patch-based MRF for a degraded word image.

The ideal (restored) patch at the location x_5 depends not only on the observed patch at y_5 , but also on the context of x_5 , given by its neighbors, x_2, x_4, x_8 , and x_6 . For example, in figure 2, the restoration of y_6 depends on whether the underlying character (b, h, n, p) , which is indicated by its neighboring patches. The goal of the restoration stage can be thought of as that of estimating the most likely set of patches, x_i , that could generate the observations, y_i ,

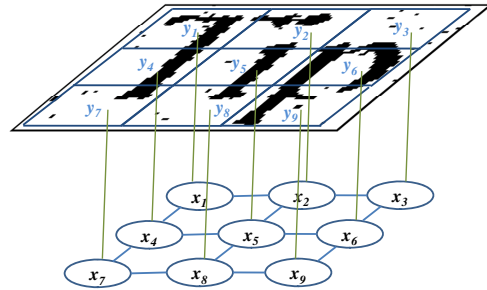


Figure 2. Patch-based MRF for a degraded word image (*Tip*).

while being in their respective neighborhoods. Based on the restoration goal, the problem in the first stage is to find a set of ideal prototypes \mathcal{X} , that are possible in a specific document, along with the probabilities of their neighborhood values, $p(x_i, x_j)$ for each of the four neighbors.

The first stage involves the estimation of ideal prototypes from degraded ones. The primary goal as mentioned before is to identify the consistent primitives (patches in our case) in the document collection. As we have a collection of documents at our disposal, we try to estimate the model from multiple observations. The process proceeds as follows: A given document image is approximately segmented into words and characters. One could make errors at this stage. The resulting segments are clustered to identify consistent shapes in the document. Errors in segmentation or highly degraded characters show up as outliers during clustering, and are eliminated from the learning phase. The consistent, probably noisy, segments are used to compute their most probable restoration. The restored segments are then covered with patches to learn their shapes and neighborhood relations.

The challenge here is to deal with the large number of possible patches at the patch size we choose, as well as to deal with the severe degradations of characters present in the document. We refer to this step as *prototype generation*. One should also be able to generate the neighborhood relationships from the prototypes. Note that using a generic MRF model as shown in figure 2 will lead to a dramatic increase in the number of possible patches. Hence the second challenge is to come up with a formulation for the restoration phase, that makes the prototype generation phase easier, and the restoration, efficient. We will first look into the restoration formulation and then return to the prototype generation phase.

2.1. Markov Model for Restoration

The input image is segmented into words, and each word is restored independently. However, we do not assume that the word segmentation is always correct. Each word is assumed to be divided into a set of possibly overlapping

patches, y_i , as shown in figure 2. Given a set of observed patches, y_i , form an input document image, I , we aim to compute the MAP (*maximum a posteriori*) estimate of the corresponding underlying labels, $x_i \in \mathcal{X}$.

Let $P(\bar{x}, \bar{y}) = P(x_1, \dots, x_N, y_1, \dots, y_N)$ be the joint probability of observing y_1, \dots, y_N when the corresponding underlying labels are x_1, \dots, x_N . Let $\psi(x_i, x_j)$ denote the pairwise compatibility of two neighboring labels, x_i and x_j , and $\phi(y_k | x_k)$, the likelihood that the label x_k generates the observed patch, y_k . The joint probability can now be written as:

$$P(\bar{x}, \bar{y}) = P(x_1, \dots, x_N, y_1, \dots, y_N) = \prod_{(i,j)} \psi(x_i, x_j) \prod_k \phi(y_k | x_k), \quad (1)$$

The first product is over all neighboring pairs of nodes, i and j . To compute the MAP estimate, we solve the MRF using the belief propagation framework [17, 5]. The belief-propagation algorithm updates *messages*, m_{ij} , from node i to node j , which are used to infer the state at node j . The state of a node is updated based on the messages it receives, and the process is repeated until convergence. Let m_{kj}^t be the message being sent from the node k to j at time instant t . The MAP estimate at node j over all label candidates x_j is:

$$\hat{x}_{jMAP} = \operatorname{argmax}_{x_j} \phi(x_j, y_j) \prod_k m_{kj}^t, \quad (2)$$

where k runs over all neighbors of node j . We calculate m_{kj}^t as

$$\begin{aligned} m_{j\uparrow}^t &= \max_{[x_k]} \vec{\psi}(x_j, x_k) \phi(x_k, y_k) m_{k\rightarrow}^{t-1} m_{k\uparrow}^{t-1} m_{k\leftarrow}^{t-1}, \\ m_{j\leftarrow}^t &= \max_{[x_k]} \vec{\psi}(x_j, x_k) \phi(x_k, y_k) m_{k\uparrow}^{t-1} m_{k\leftarrow}^{t-1} m_{k\downarrow}^{t-1}, \\ m_{j\downarrow}^t &= \max_{[x_k]} \vec{\psi}(x_j, x_k) \phi(x_k, y_k) m_{k\leftarrow}^{t-1} m_{k\downarrow}^{t-1} m_{k\rightarrow}^{t-1}, \\ m_{j\rightarrow}^t &= \max_{[x_k]} \vec{\psi}(x_j, x_k) \phi(x_k, y_k) m_{k\downarrow}^{t-1} m_{k\rightarrow}^{t-1} m_{k\uparrow}^{t-1}. \end{aligned} \quad (3)$$

m_{kj}^{t-1} is m_{lk}^t from the previous iteration. The initial m_{kj}^0 s are set to column vectors of 1's, of the dimensionality of the variable x_j . Spatial constraints are imposed through the formulation of $\vec{\psi}(x_j, x_k)$ function. Here, ψ is not a symmetric function and depends on the orientation of x_j and x_k , enabling the prior being stronger than smoothness prior [21].

2.2. Localizing the Patches

As noted before, one of the main constraints in the patch based formulation is that the location of structures within a patch can vary, changing the observation probabilities,

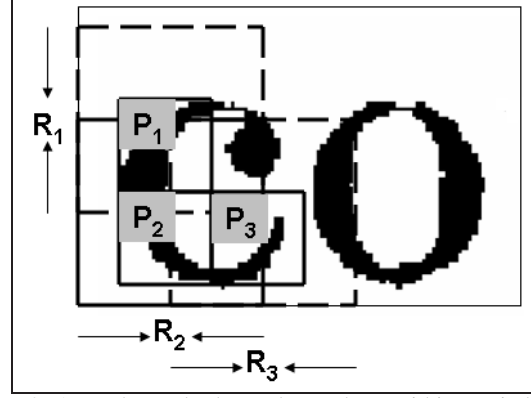


Figure 3. A patch can be located anywhere within a window of $m \times n$ within the word image.

$\phi(x_k, y_k)$. To deal with this, we allow the patches to slide around and settle at a location that best matches the underlying label. The spirit of our approach is similar to [2]. However, we note that the label itself is unknown and further depends on its neighboring nodes. Hence we need to carry out the optimization procedure described above for all possible patch locations for each of the patches. Let each patch, y_k be offset in the horizontal and vertical directions by Δp_k and Δq_k respectively, from their initial uniformly spaced locations. The function to optimize becomes:

$$P(\bar{x}, \bar{y}) = \max_{\Delta p_k, \Delta q_k} \prod_{(i,j)} \psi(x_i, x_j) \prod_k \phi(y_k, \delta p_k, \delta q_k | x_k) \quad (4)$$

The direct optimization of equation 4 over all patch locations turns out to be prohibitively expensive. To overcome the difficulty, we enforce left-to-right and top-to-bottom orderings on the centroids of the patches and formulate a dynamic programming solution to carry out the computation. We further observe that the vertical sliding of a patch within a word is limited and hence for each horizontal position for a patch, we compute the best matching vertical position for every interpretation of the underlying label.

The problem is now reduced to finding the best horizontal shift for each patch. To achieve this, we apply a Viterbi decoder for every row of nodes, while keeping the patches in other rows, anchored to their locations. The process is repeated, sequentially, until convergence. Note that we need to carry out an MRF optimization at every step in the Viterbi algorithm. The initialization of the patches is carried out using independent maximum likelihood estimates for the patches over all possible labels and locations within the limits. We can further improve the computation speed by restricting the range of sliding for each patch to a specific limit, restricting the most likely path (horizontal locations) within a diagonal band, referred to as the Sakoe-Chiba band [18].

A lighter version of the optimization can be obtained if



Figure 4. Collection of Characters and their Prototypes. A collection of 10 characters are used to generate the prototype.

we assume that the position of a patch is within one window width around its initial location. This makes the computation of path locations independent of its neighbors, and the resulting optimization function would be:

$$P(\bar{x}, \bar{y}) = \prod_{(i,j)} \psi(x_i, x_j) \prod_k \max_{\delta p, \delta q} \phi(y_{k, \Delta p, \Delta q} | x_k) \quad (5)$$

$$= \prod_{(i,j)} \psi(x_i, x_j) \prod_k \phi'(y_k | x_k) \quad (6)$$

Note that the above equation leads to a regular MRF formulation. In most practical cases, we found that the direct MRF formulation using equation 6 leads to the same solution as the more complex Viterbi optimization using equation 4.

3. Learning the Labels and Context

To generate the label set, we collect a set of similar characters by segmentation and clustering. Outliers in each cluster are usually errors in segmentation process or highly corrupted samples, and are removed [11]. These similar characters are used to generate high quality prototypes, using a MAP estimate. Similarly, prototypes are generated from all the clusters. Figure 4 shows examples of two prototypes, corresponding to characters ‘a’ and ‘d’ being generated from the noisy samples.

To learn the context relationship, each character prototype is divided into a set of patches. Each character is divided into patches of size $N \times N$, and the number of patches depend on the dimensions of the character. Similarly labels are extracted from each character prototype. The collection of patches from all the characters form the possible set of labels. These patches are typically of size 25×25 for a 600 dpi image. Large patch size means that the prior is defined on large neighborhood, making it more powerful.

We sample the patches so that they overlap with each other by few pixels. In the overlap region, the pixel values of compatible neighboring patches should agree. We measure $d(x_i, x_j)$, the sum of squared differences between patch candidates x_i and x_j in their overlap regions at nodes i and j . The compatibility matrix between nodes i and j is

then

$$\vec{\psi}(x_i, x_j) = \exp\left(-\frac{\vec{d}(x_i, x_j)}{2\sigma^2}\right), \quad (7)$$

where σ is a noise parameter [6]. We use a correlation based penalty on differences between the observed degraded image patch, y_i , and the candidate label patch, x_i , found from the prototype to specify $\phi(y_k | x_k)$.

3.1. Document Image Super-resolution

One of the advantages of our formulation of learning ideal patches from multiple degraded or low-resolution patches is that we can directly estimate the ideal patches at high resolution, thus combining document restoration and super-resolution into one process. We now look into the use of a MRF-based MAP estimation to generate the super-resolved prototypes. The overall process proceeds as follows:

- Upsample the low-resolution, degraded prototypes using cubic spline interpolation.
- Register the prototypes at high resolution using correlation.
- Obtain the super-resolved patches by computing the MAP estimate of the underlying high resolution prototype.

To achieve the third step, we use a text specific prior and formulate the estimation in an MRF framework. Images of text are usually smooth in both the foreground and background, with sharp transitions only at the edges. Thus, text images typically have bimodal distributions, with large black and white peaks [4]. The peak occurs at μ_{white} for the background (white), since the majority of pixels on a text page is background. There is a second peak at μ_{black} , representing the text. Additionally, there are a small number of gray values occurring between the two peaks, which represent the gray pixels that exist at transitions from white to black. The amount of these intermediate gray levels is related to the amount of blur in the document image. In order to obtain an unblurred image, we wish to obtain a sharp bimodal distribution, pushing the intermediate gray level towards their nearest peaks. To incorporate this property we define the conditional probability as

$$\zeta(y_k | F) = (y_k - \mu_{white})^2 (y_k - \mu_{black})^2, \quad (8)$$

where $\zeta(y_k | F)$ is the cost of assigning label x_k to pixel y_k , affecting the (bimodal) distribution F , and is referred to as the bimodal cost prior.

We would also like the label x_k to be as close to the gray value y_k , for a pixel. Thus, the conditional probability is

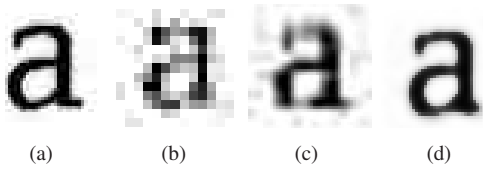


Figure 5. Super-resolution by a factor of 3: (a) original high-resolution image, (b) low-resolution input (c) cubic spline interpolation of (b), and (d) super-resolved prototype.

defined as:

$$\phi(y_k|x_k) = (y_k - x_k)^2 \zeta(y_k|F) \quad (9)$$

where $\phi(y_k|x_k)$ is the cost of assigning label x_k to pixel y_k , which is referred to as the data cost. We use an edge stopping function to ensure sharp edges. Thus we use Lorentzian edge penalty function [15], which determines the penalty between the two nodes of a MRF:

$$\psi(x_i, x_j) = \log \left\{ 1 + \frac{1}{2} \frac{(x_i - x_j)^2}{\sigma_L} \right\}, \quad (10)$$

where σ_L is called the contrast parameter, which controls the shape of the edge stopping function [15]. The quality of a labeling in general restoration problem is given by an probability estimate in equation 1. Thus, unlike Luong et al. [15], we formulate the problem as MRF that provides us with a better optimum. The belief propagation [5] based optimization is both fast and robust for the purpose.

4. Experimental Results and Discussions

The experiments were designed to help us to analyze the performance of the algorithms as well as give insights into its working and potential applications. We now discuss some of the quantitative and qualitative results on various input documents.

4.1. Restoration of Degraded Documents

We have carried out a variety of restoration experiments with different document images and differing levels of degradation. For the first experiment, we selected a degraded English book containing 40 pages with close to 50,000 words and 237,000 characters. The pages of the book were scanned using an HP flatbed scanner at 600dpi. A document model was learned for the complete book after segmentation, and restoration was performed on all the pages. Figure 7 shows a selection of 10 words from the book containing cuts, merges, blobs and erosion artifacts, along with the restoration output of our algorithm.

The first class of degradations that we notice is ink blobs and smears, as present in the word *surely*, *convening*, *permitting*, etc. Our algorithm is able to handle most of the

cases very well. Especially, the word *little*, which had three of its characters connected by an ink smear was restored correctly. Sever and minor cuts and erosions were also present in the dataset. For example, the word *several* has a severe cut in the character *v* and the character *m* in *imprisonment* is separated into three parts due to erosion. As the overall shape of all the characters are discernible in spite of these degradations, the restoration algorithm is able to replace the noisy regions with the correct ideal patches.

However, we note that for the word *vindictive*, the ink smear on the character *v* is so severe that the resultant patches were not correctly matched. As the restoration always tries to find a set of patches whose neighborhood relations are correct as per the document model, we notice the patch replacements have resulted in replacements by patches of character *a*.

The restoration should also improve the recognition results of any off-the-shelf OCR system. To verify this, we ran the Tesseract-2.01 OCR from Google on all the pages of the above book, which resulted in an error rate of 3.7%. We note that the modern day OCRs are trained to perform well even in presence of common types of noise, and the accuracy on the original document is already very good. However, after restoration by our proposed algorithm, the error rate further reduced to 1.9%. The accuracy was measured at character level and the book contained 236,861 characters.

Figure 8 shows a portion of an input page and the restored version, along with the OCR results. We note that the recognition of the restored document is in fact highly accurate, and most of the errors are introduced during the rectification process. Two types of errors are of interest here. The first one is due to the erroneous restoration of the word *vindictive*, where the ink-blotted *v* was restored as

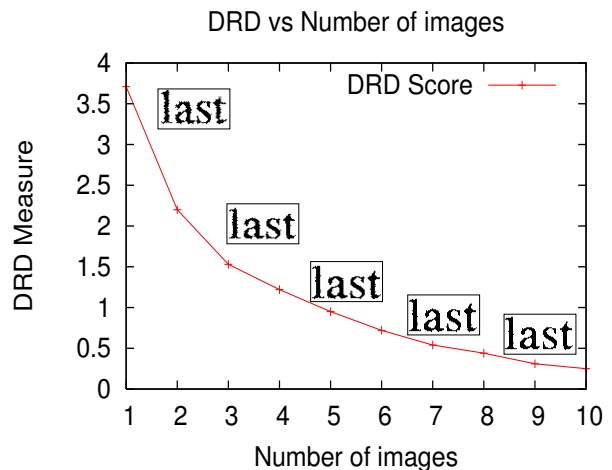


Figure 6. Distance-Reciprocal Distortion Measure for a word “last”.

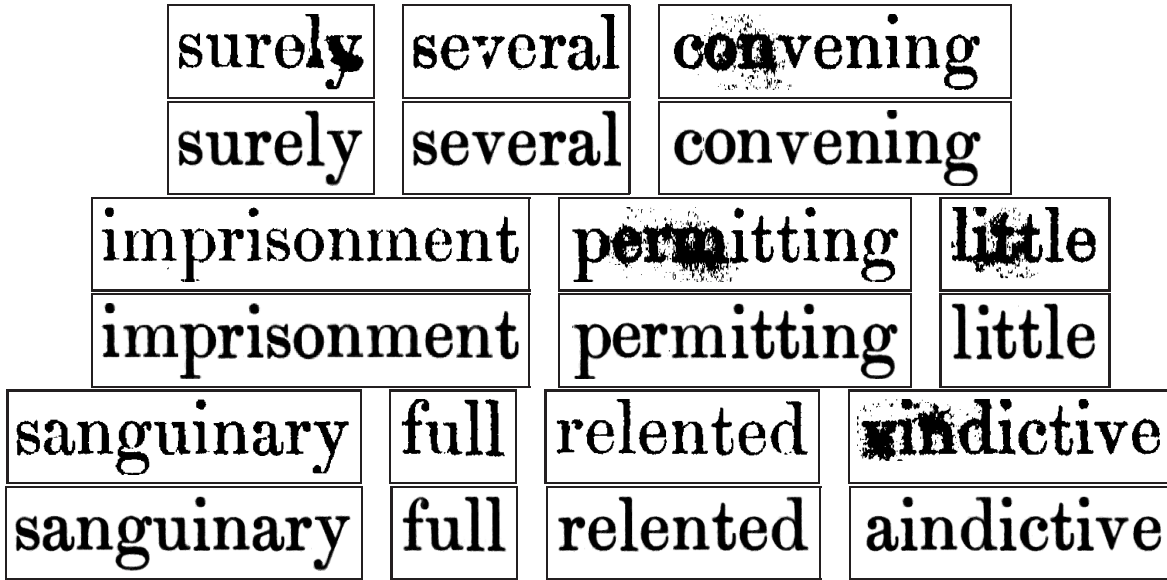


Figure 7. Restoration of words containing cuts, merges, blobs and erosion.

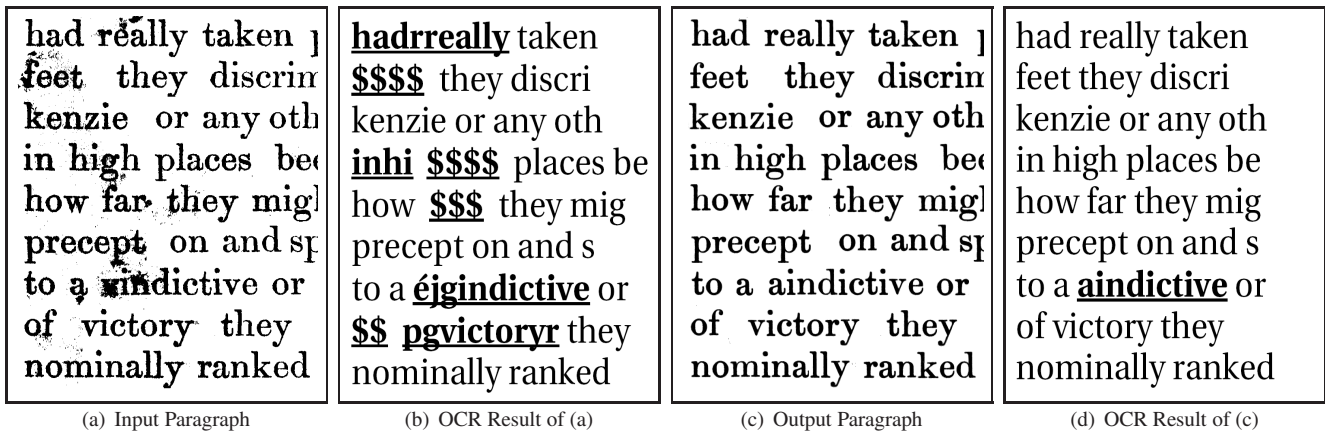


Figure 8. Result on a portion of image from the book.

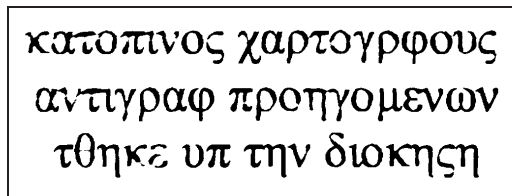
an *a*. The second set of errors is due to missing punctuation marks. This is primarily because of the assumption of heavy noise in documents during prototype learning, which discards small marks. One can tune the restoration to the noise levels present in the document to avoid this.

To study the effect of the size of the document on the restoration results, we analyze the restoration quality with increasing number of prototypes available in each cluster. To measure the restoration quality, we use *distance-reciprocal distance measure* [14], defined as: $DRD = (\sum_{k=1}^S DRD_k) / NUBN$, where *NUBN* estimates the nonempty area in the image and DRD_k is the weighted sum of the pixels in the block of the original image that differ from the flipped pixel in the degraded image. We select one word from the above book and plot the DRD score as the

number of prototypes used in the learning stage increases. We note that with around 7 prototypes, the *DRD* score is already very low, which keeps improving further over iterations. We also show a sample restored word if performed at different stages of the learning process for illustration purposes. One can clearly notice the increase in visual quality of the sample word as the number of prototypes in a cluster increases.

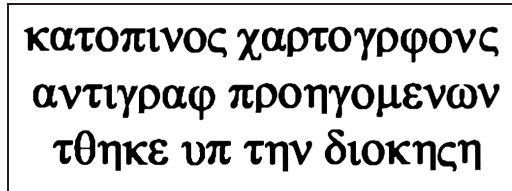
4.2. Script Independence

As mentioned before, the restoration process does not make any assumption about the font or script of the input document. We can perform restorations of multiple scripts using the same approach. As the learning happens at a patch level, and the document priors are generic to text, the algo-



κατοπινος χαρτογρφους
αντιγραφ προηγομενων
τθηκε υπ την διοκηση

(a) Input document



κατοπινος χαρτογρφους
αντιγραφ προηγομενων
τθηκε υπ την διοκηση

(b) Restored document

Figure 9. Restoration of text in Greek using the proposed approach.

rithm is directly applicable to any script or font. The result of restoration of the proposed algorithm on a document containing Greek text is shown in figure 9. Several touching and broken characters are effectively corrected by our restoration algorithm.

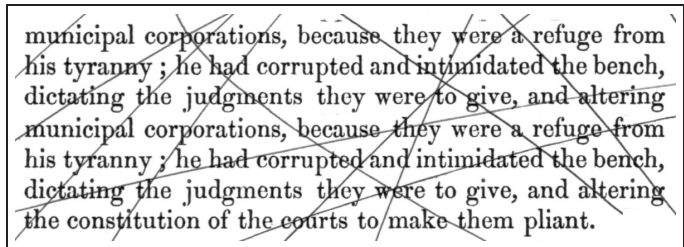
4.3. Restoration of Vandalized Documents

One of the main strength of the approach is that it models the contents of the document image. This allows us to discard any additions to the document that do not follow the learned document model. We are able to restore even severely degraded and vandalized documents, as long as the actual content is discernible. As the learning is done at a patch level, one can learn the document model from the degraded/vandalized document itself, assuming parts of the document has segmentable patches that can be used for learning. Figure 10 shows an example document that has severe scratches/overwriting on the original document. We notice that our approach is able to completely recover the original document. Restoration results of document images from a magazine with degradations and vandalization are given in figure 11.

It is interesting to note that even in presence of severe degradations, our algorithm is able to perform extremely well, as long as the overall shape of a patches in a character or its neighbors are visible.

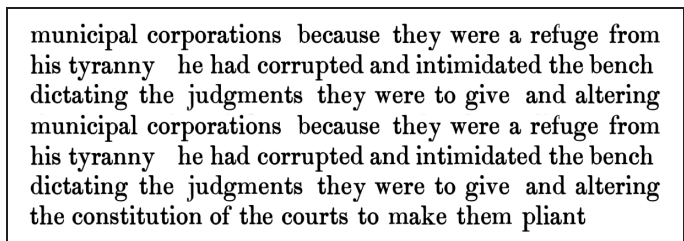
4.4. Restoration with Super-resolution

Another aspect of the approach of using multiple degraded prototypes to learn the ideal one is that one can infer super-resolved prototypes for patch models and restoration. Figure 12 shows a sample document at 100dpi that is super-resolved to 300dpi. Comparison with the original document scanned at 300 dpi reveals that while the super-resolved text



municipal corporations, because they were a refuge from his tyranny ; he had corrupted and intimidated the bench, dictating the judgments they were to give, and altering municipal corporations, because they were a refuge from his tyranny ; he had corrupted and intimidated the bench, dictating the judgments they were to give, and altering the constitution of the courts to make them pliant.

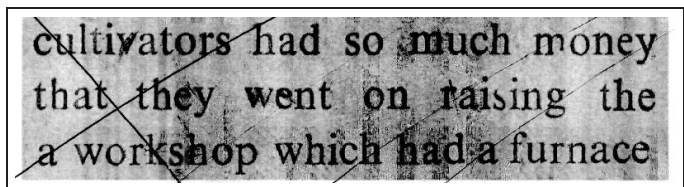
(a) Vandalized input document



municipal corporations because they were a refuge from his tyranny he had corrupted and intimidated the bench dictating the judgments they were to give and altering municipal corporations because they were a refuge from his tyranny he had corrupted and intimidated the bench dictating the judgments they were to give and altering the constitution of the courts to make them pliant

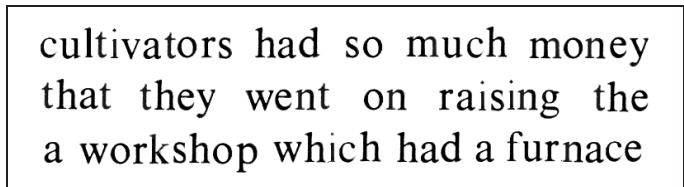
(b) Restored document

Figure 10. Restoration results of a page with overwritten scratches.



cultivators had so much money that they went on raising the a workshop which had a furnace

(a) Portion of a magazine page



cultivators had so much money that they went on raising the a workshop which had a furnace

(b) Restored page

Figure 11. Restoration results of document images from a magazine.

is close to the original, the process has achieved its intended goal of restoration (noise removal) also.

5. Conclusions and Future Work

We presented a novel approach to document restoration, that builds a tight model of the document content from the input document itself, and uses it to restore severe degradations, including cuts, merges, blobs and erosions. Modeling the document as an MRF on larger patches allows us to use a larger context for restoration. As the approach works on a generic model of the document content, we are able to handle vandalized documents as well as multiple scripts and fonts. The estimation of the content model can also

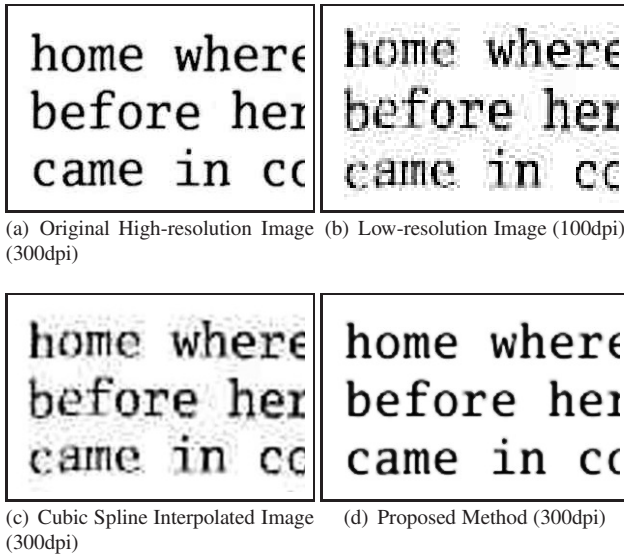


Figure 12. Text super-resolved by a factor of 3 times.

incorporate generation of high quality prototypes, leading to super-resolution of the restored document. The quality of restoration is especially high for larger document collections such as pages from a book, and hence is ideal for large document collections such as digital libraries.

The current approach primarily uses a content model that is learned from the input document. Integration of the approach with a complementary mechanism that models the nature of degradations could further improve the restoration performance. Another potential direction is to combine recognition with restoration in an iterative fashion.

References

- [1] Digital library of India. <http://dli.iit.ac.in/>.
- [2] E. Borenstein and S. Ullman. Combined top-down/bottom-up segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(12):2109–2125, 2008.
- [3] H. Cao and V. Govindaraju. Handwritten carbon form pre-processing based on markov random field. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
- [4] K. Donaldson and G. K. Myers. Bayesian super-resolution of text in video with a text-specific bimodal prior. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [5] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient belief propagation for early vision. *International Journal of Computer Vision*, 70(1):41–54, 2006.
- [6] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Comput. Graph. Appl.*, 22(2):56–65, 2002.
- [7] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 20(1):25–47, 2000.
- [8] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.
- [9] R. Gonzalez and R. Woods. *Digital Image Processing*. Prentice Hall, 2002.
- [10] M. D. Gupta, S. Rajaram, N. Petrovic, and T. S. Huang. Restoration and recognition in a loop. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [11] J. D. Hobby and T. K. Ho. Enhancing degraded document images via bitmap clustering and averaging. In *Proceedings of the International Conference on Document Analysis and Recognition*, 1997.
- [12] Y. Huang, M. S. Brown, and D. Xu. A framework for reducing ink-bleed in old documents. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.
- [13] S. Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995.
- [14] H. Lu, A. Kot, and Y. Shi. Distance-reciprocal distortion measure for binary document images. *IEEE Signal Processing Letters*, 11(2):228–231, 2004.
- [15] H. Q. Luong and W. Philips. Robust reconstruction of low-resolution document images by exploiting repetitive character behaviour. *International Journal of Document Analysis and Recognition*, 11(1):39–51, 2008.
- [16] G. Myers and K. Donaldson. Bayesian super-resolution of text in video with a text-specific bimodal prior. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1188–1195, 2005.
- [17] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann, 1988.
- [18] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Speech and Audio Processing*, 26:43–49, 1978.
- [19] P. K. Sankar, V. Ambati, L. Hari, and C. Jawahar. Digitizing a million books: Challenges for document analysis. In *Proceedings of DAS-2006*, pages 425–436, 2006.
- [20] P. Sarkar, H. Baird, and X. Zhang. Training on severely degraded text-line images. In *Proceedings of the International Conference on Document Analysis and Recognition*, pages 38–43, 2003.
- [21] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. F. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 30(6):1068–1080, 2008.
- [22] Q. Wang, T. Xia, L. Li, and C. L. Tan. Document image enhancement using directional wavelet. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003.