

Inpainting, y otras tecnicas de reconstruccion de imagenes aplicado en documentos danados o censurados

Jahir Medina

Miguel Pacheco

Resumen

Inpainting (*pintando dentro de*), es una tecnica de reconstruccion de imagenes, usado para reconstruir informacion perdida o danada en la imagen. Algunas tecnicas reconstruccion, que pueden estar basadas en operaciones morfologicas y variantes de otros algoritmos seran usados para reconstruir informacion en imagenes de documentos, los cuales podrian estar danados o censurados (tachas de corrector, o manchas de tinta).

Para favorecer esta reconstruccion, se emplea una fuente de luz intensa, de manera que el documento se encuentre a contraluz; de luz blanca o de distintos colores.

Introduccion

Realidad Problematica

Muchos documentos, que durante el proceso de edicion tuvieron erratas y fueron posteriormente editados pueden llegar a poseer danos, aun que voluntarios, excesivos.

Si bien parte de esto puede ser reconstruido reescribiendo el documento, para casos donde el documento es antiguo (considerese registro de notas de los años 1970, en papel bulki y escrito a maquina) y no solo debe ser realizado en solo un documento, resulta excesivo y lento.

En grandes ficheros, donde una solucion moderna puede ser escanear documentos y aplicar un post procesamiento con algoritmos OCR, muchos elementos son omitidos:

- imagenes

o en 12 créditos en
re y cuando el estudia
l quedan exonerados del
re-requisitos académicos
o, ~~se~~ matrícula de estos
condición.

Figure 1: Ejemplo de Correcciones

- tablas
- caracteres blanqueados
- borraduras
- destinturación de caracteres

Esto disminuye significativamente la automatización del proceso bajo técnicas estándares, además de que implementar técnicas modernas basadas en aprendizaje automático resultaría en el rediseño de hardware ya adquirido.

Alternativa de Solución

- Emplear 4 luces diferentes : Blanco, Rojo, Azul, Verde.
- Usar dichas luces como fuente de luz trasera al documento

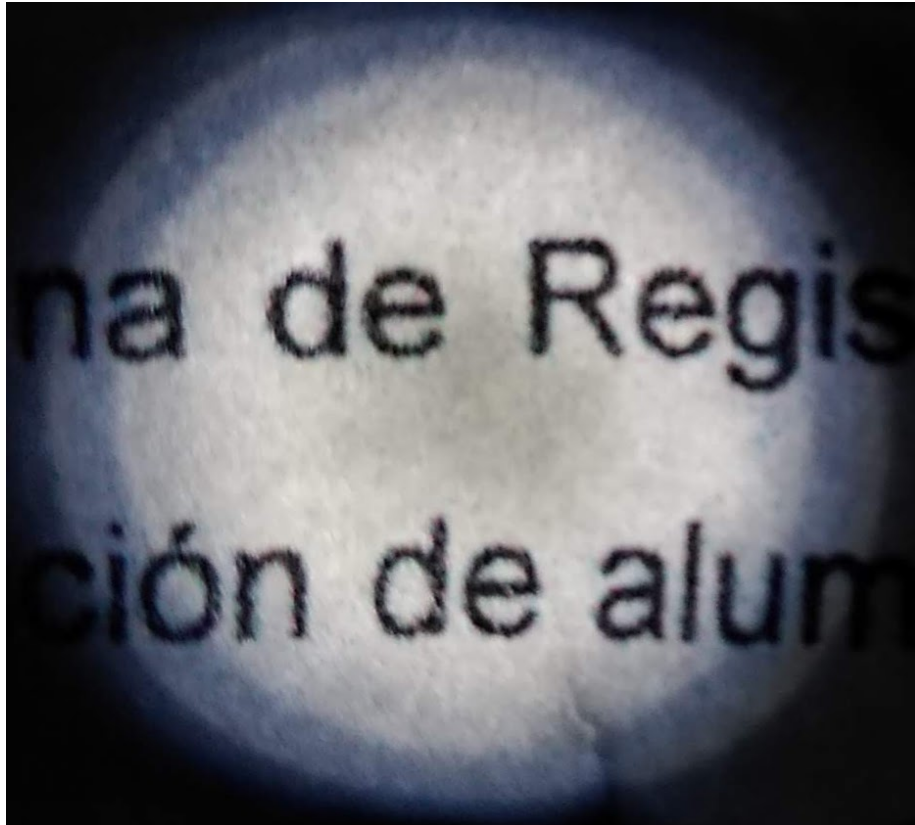


Figure 2:

- Capturar imágenes del documento y las partes dañadas específicas
- Reconstruir las partes dañadas específicas mediante técnicas clásicas.

- reduccion de ruido
- operaciones morfologicas (eliminacion de protuberancias)
- separacion de figura y fondo
- segmentacion
- inpainting (reconstruccion de las partes faltantes)
- integracion al contexto general
- Puesta en “disponible” para su postprocesamiento mediante algoritmos OCR.

Algunos ejemplos de censuras:

Tinta (Lapciero)



Figure 3: Tinta (Lapciero)

Corrector (Blanco)

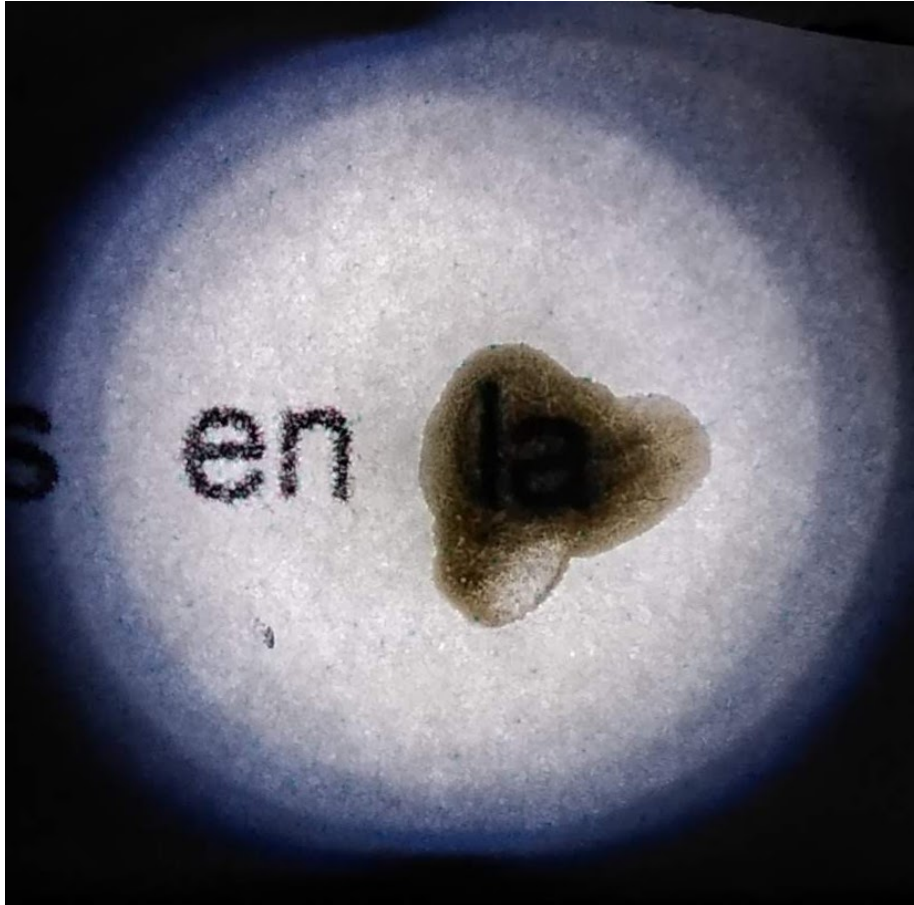


Figure 4: Corrector (Blanco)

Tinta sobre Corrector

Marco Teorico

Placeholder

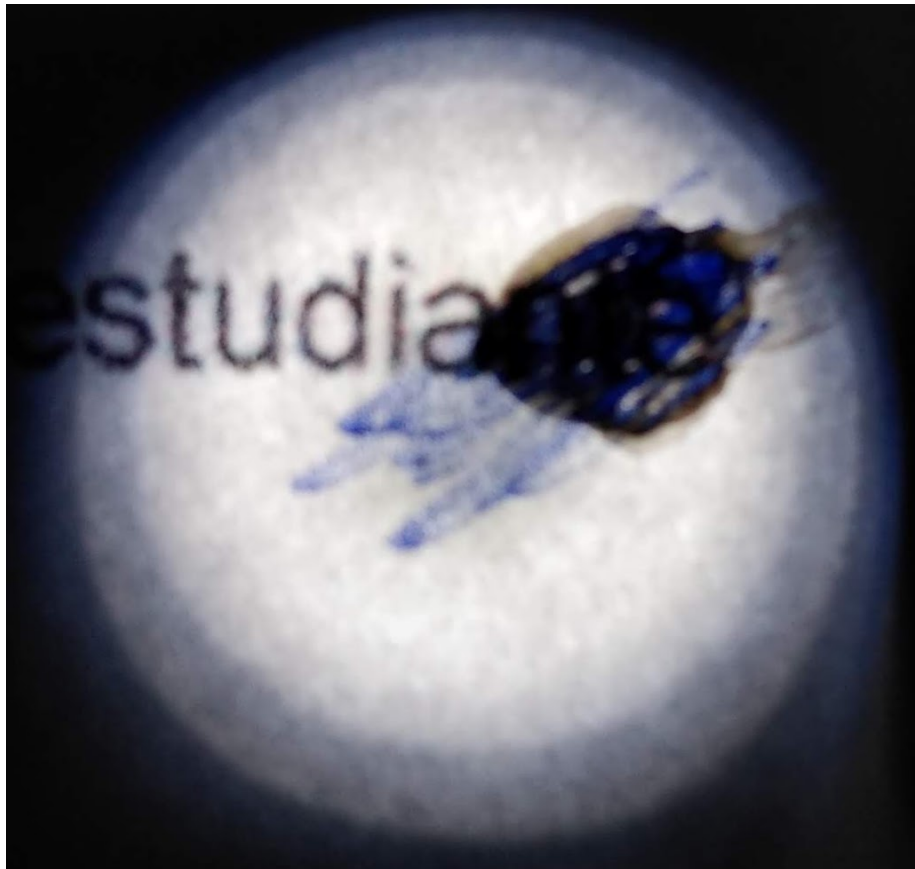


Figure 5: Tinta sobre Corrector

Metodologia

Placeholder