

extract "paper" from document scan?

Source: <http://sci.tech-archive.net/Archive/sci.image.processing/2008-04/msg00013.html>

- *From:* bugbear <bugbear@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Wed, 02 Apr 2008 10:21:04 +0100
-

I'm scanning old books, maps, pamphlets etc.

In general, I have either uneven lighting, faded or stained originals, or both.

What I would like to do is extract "Just the paper", and then smooth (blur?) it, and then use this as a lighting (and/or staining) reference for processing the main scan.

At the hand waving level, I would like to remove all the marks (defined as "darker than some threshold value") and then replace the pixels I just removed with pixels generated by interpolation from the remaining pixels.

But I can't quite put my finger on a sequence that works.

I would (ideally) like to code this as a shell (or perl) script driving netpbm or imagemagick (I'm working on Linux), but all suggestions will be given careful attention.

Thanks for any help you can give me,

BugBear

.