

Auto Multiple Choice - Feature # 66: Use multiple pages pdf as scanned answered sheets

Status:	Closed	Priority:	Normal
Author:	Nicolas Pettiaux	Category:	
Created:	06/26/2012	Assignee:	
Updated:	09/22/2012	Due date:	
Description:	The scanner I use produces multiple pages pdf as the results of scanning the answered sheets. Today I need to convert the pdf (how to do easily with a script ?) to good quality jpeg manually with gimp (as I do not know how to script gimp nor have found the parameters to have Imagemagick convert to it as well as gimp).		

History

06/27/2012 07:47 am - red sea

you can use these programs:

- pdf chain
- pdf sam

regards

06/27/2012 08:31 am - Alexis Bienvenüe

Another suggestion: as PDF is only a container, extracting the included images can be done without modifying them using @pdfimages@ from the @poppler-utils@ debian package:

```
<pre>pdfimages scan.pdf scanroot
</pre>
```

In some situations however, the resulting images will need 90° rotation...

06/27/2012 08:50 am - Nicolas Pettiaux

Alexis Bienvenüe wrote:

> Another suggestion: as PDF is only a container, extracting the included images can be done without modifying them using @pdfimages@ from the @poppler-utils@ debian package:

> [...]

> In some situations however, the resulting images will need 90° rotation...

Great.

In my case, the resulting file is in format im-001-000.pbm: Netpbm PBM "rawbits" image data. Does AMC accept this format ? or do I have to convert it to jpeg ? or another ? which is the best way and with which parameters ?

06/27/2012 08:54 am - Alexis Bienvenüe

> Netpbm PBM "rawbits" image data. Does AMC accept this format ?

This should be OK, as this format can be read by OpenCV.

06/27/2012 09:16 am - Nicolas Pettiaux

good to know. Problem is solved then. Much thanks.

07/07/2012 05:47 pm - Alexis Bienvenüe

Also note that AMC accepts multi-page PDF scans as input when calling "automatic data capture": it does convert it to bitmap single page images before processing them.

07/08/2012 09:57 am - Nicolas Pettiaux

Alexis Bienvenüe wrote:

> Also note that AMC accepts multi-page PDF scans as input when calling "automatic data capture": it does convert it to bitmap single page images

before processing them.

Thanks for the info. So, if I understand well, the situation now is : we can use with AMC a multipage pdf scan, that is created by the copy machine acting as scanner. The rest will be automatic ?

07/08/2012 11:12 am - Alexis Bienvenüe

Yes.

09/06/2012 09:32 am - Alexis Bienvenüe

- *Status changed from New to Resolved*

09/22/2012 10:38 pm - Alexis Bienvenüe

- *% Done changed from 0 to 100*

- *Status changed from Resolved to Closed*