

スキャン補助用Photoshopスクリプト : CS-PS4CropPicture

Copyright (C) 2011 つぶ@coffee shop (tsubu@coffee shop)

できること

スキャンした画像に対し以下のpost-processingが行えます。

- ICCプロファイルの指定
- 背景の除去 (remove the black border) : ある程度の割切りができる人向け

色調整等の前処理にでも使ってください。

必要要件

本プログラムはPhotoshopのスクリプトのため、実行するにはAdobe Photoshopが別途必要となります。

テストはWindows7 (64bit)、Photoshop CS4 (64bit版) という環境で実施しています。

他のバージョンでも動くと思いますが、論理的な要件として以下があります。

- スクリプトに対応したPhotoshop (デフォルトで対応しているのはCS以降)
- ScanTailorのプロジェクト対応のため CS3以降のPhotoshop

留意・注意事項

- 本ソフトウェアはベータ版ということもあり**再配布は許諾していません**。現在の配布場所はベクター様になります。

(Author **don't permit the re-distribution** of this script(s) because of some reasons.)

- 著作権者に無断で拡散を行うための作業(俗にいう放流目的)に、本スクリプトの使用は許諾しません。
- 本ソフトウェアは、無保証であり全て使用された方の責任においてご利用ください。

使用に起因するいかなる事態も責任を負いかねますので、事前に十分にテストをし問題がないことを確かめてから本番で使用して下さい。

(ぶっちゃけ商業作品と比べるとテストが不十分です。パラメータにはまともな値を渡してください。

不特定多数に公開するような形態ではなく、かつ自分のための作業にあえて変な値を渡すこともなかろうと判断し、チェックをさぼっています。)

本ソフトウェアは上記、留意・注意事項を順守することを条件に無償で利用できます。

無償で利用できますが、百回、千回ぐらいぶんまわし時間が節約ができたなど、ありがたみを感じて頂けた場合にご寄付頂ければ、開発環境を維持する上で非常に助かります。

インストール

GUI用のスクリプト CS-PS4CropPicture_GUI.jsx とバッチ処理を行うスクリプト CS-PS4CropPicture_process.jsxの2部構成になっています。

Windowsで拡張子を関連付けている場合(Photoshopのデフォルトのはず)は、そのままダブルクリックしてすぐに使うことができます。

Photoshopのメニューに出させるなどファイルを移動する場合は、同梱のスクリプトは同じ場所に配置して下さい。

蛇足ですが、Windowsの64bit版では32bitもインストールされているので、そちらに入れてメニューに出ないと無駄に時間を費やすことのないようにしましょう。(明日の方向をみながら)

アンインストール

単なるスクリプトですのでファイルを削除するだけでアンインストールできます。

ソフトウェアの使い方

本スクリプトはスキャナで取り込んだ画像のバッチ処理を想定しています。

1. ソースとする画像ファイルを特定の構成で配置した後、
2. スクリプトを実行して下さい。

スクリプトを実行すると、加工を行った画像ファイルを出力フォルダへ吐き出します。

ソースとなる画像ファイルには手を加えません。出力フォルダへ別名保存しています。

ソースとする画像について

スクリプトが取り扱う画像ファイルはTIFF, PNG, JPEG, BMPのカラー、グレースケールで、2値画像は対象にしていません。

出力されるファイルは入力ファイルに応じたファイル形式でなるべく保存するようにしています。

入力フォーマット		出力フォーマット	補足
TIFF	→	TIFF (無圧縮)	
JPEG	→	TIFF (無圧縮)	
BMP	→	BMP	16bit/channelでは保存できません
PNG	→	PNG (非インターレース)	保存が非常に遅かった

ICCプロファイルの指定は、自分でプロファイルを作成していたり、メーカーのプロファイルを指定したい場合に指定して下さい。

指定するプロファイルがメニューにない場合はスクリプトに追加して下さい。

ICCプロファイルがよく分からない場合はかえっておかしくなるので、この機能は使わない方がよいと思います。

背景除去機能を利用する場合の留意事項

- スキャン画像はL判などある程度の大きさがあることを前提にしています。
- 原稿部分に対し背景部分が大きければ大きいほど余計に処理時間がかかります。
- また背景部分が、中の原稿部分に比べ大きすぎる場合うまく動作しません。背景が原稿の2倍以上は事前に切りつめて下さい。

- 背景部分と原稿部分にある程度色に差があることを前提としています。背景色とほとんど差がない場合、本スクリプトではうまく動作しません。2階調化で背景と分離可能か確認し、分離できるようならば使用して下さい。

ファイル構成 (実行前に用意)

スクリプトを実行する前に処理するファイルを次のようなフォルダ構成に配置して下さい。

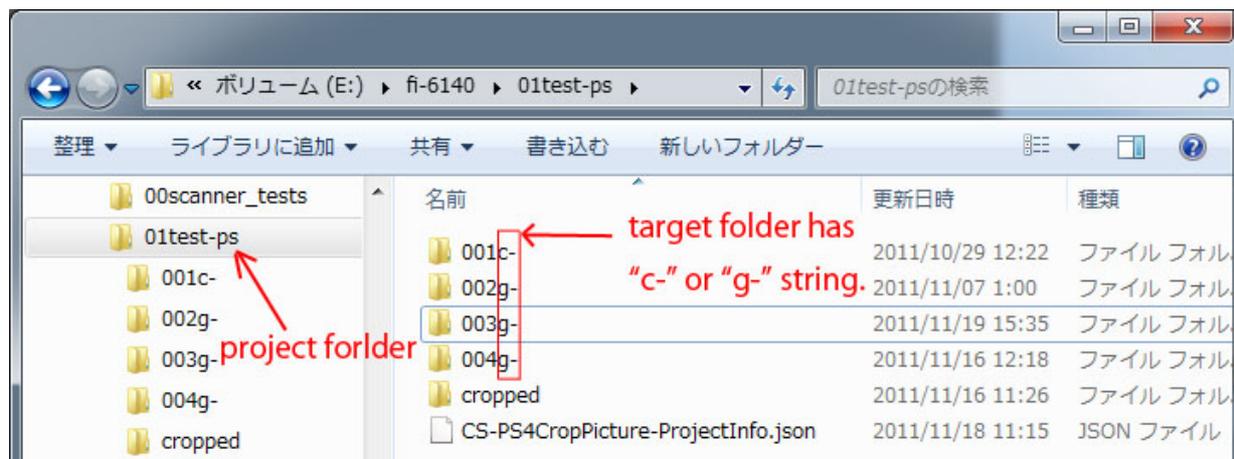
起点となるフォルダをプロジェクトフォルダと呼ぶことにすると、プロジェクトフォルダの下に001c-、002g-、……といった感じでフォルダを作成して下さい。ここではそれらを処理対象フォルダと呼ぶことにします。

フォルダ名に "c-" がついている処理対象フォルダは、カラー原稿が格納されているとみなして配下のファイルを処理します。具体的には下の処理を行います。

- カラーのICCプロファイルを指定していた場合、そのプロファイルを指定します。

フォルダ名に "g-" がついている処理対象フォルダは、グレースケール原稿が格納されているとみなして配下のファイルを処理します。具体的には下の処理を行います。

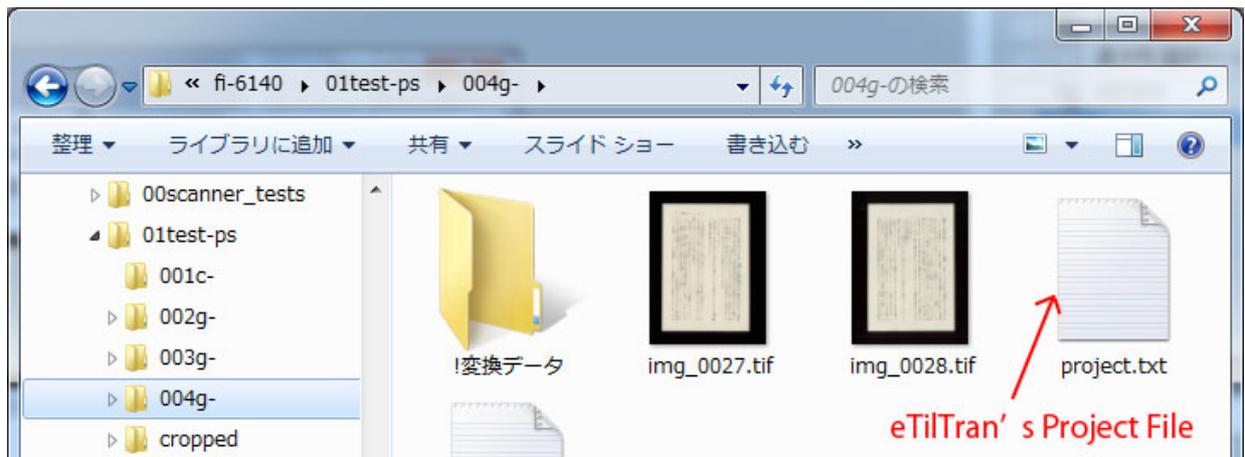
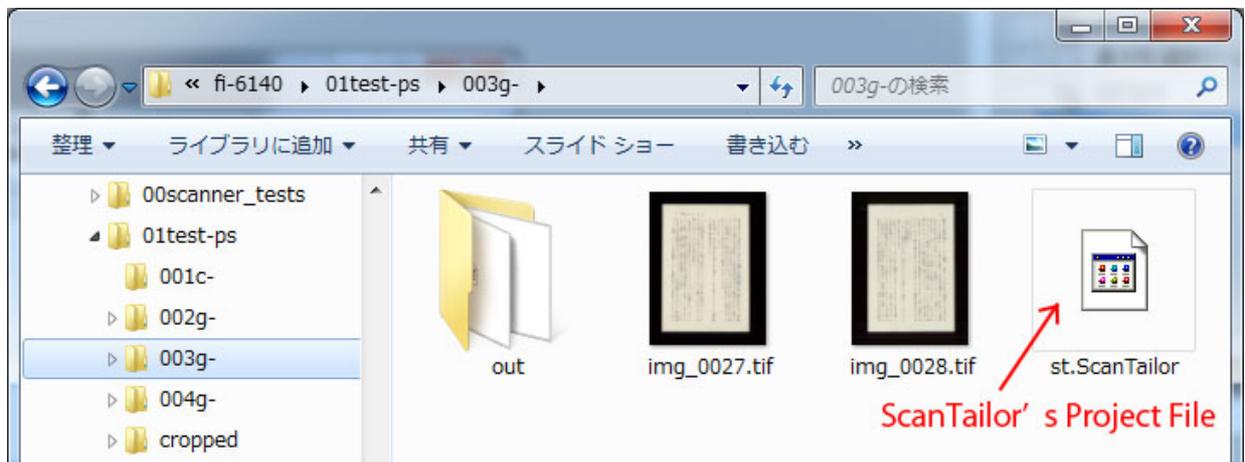
- 画像がグレースケール画像でない場合、カラーからグレースケールへ変換します。
- グレーのICCプロファイルを指定していた場合、そのプロファイルを指定します。



蛇足ですがフォルダ名は001c-o, 001c-e という感じの名称にしても処理可能です。(oはodd(奇数)、eはeven(偶数)の意味)

次に、処理対象フォルダの下にスキャン画像を置いて下さい。事前のファイル配置で必要なことはこれだけです。

もし角度補正もさせたい場合、Scan TailorないしeTilTranでプロジェクトファイルを作成し、処理対象フォルダの下に置いて下さい。Scan TailorやeTilTranのプロジェクトファイルが処理対象フォルダにあれば、処理を行う際にプロジェクトファイルから情報を取り出し画像を回転します。



プロジェクトファイルを参考にしたバージョンは次になります。これらのアプリケーションで今後プロジェクトファイルのフォーマットが変更された場合は使用できなくなる可能性が高いです。

アプリケーション名	バージョン番号
Scan Tailor	0.9.11.1
eTilTran	V0.12β9

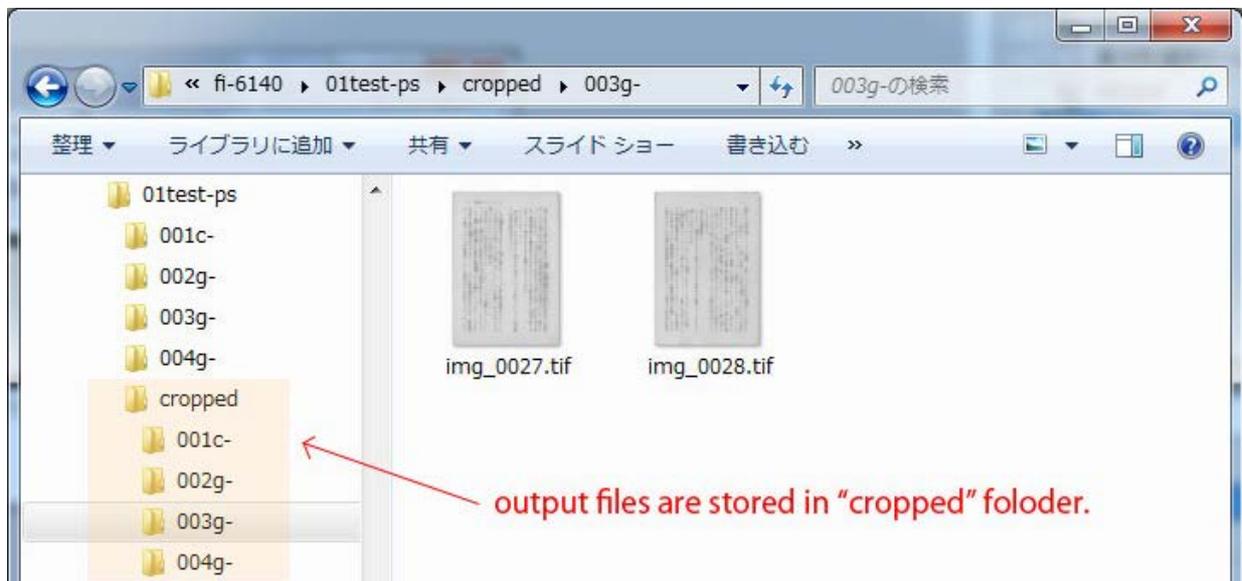
申し訳ないのですがプロジェクトファイルの名称は現状決め打ちで、Scan Tailorの場合 st.ScanTailor、eTilTranの場合 project.txt (defaultの名称) とさせて下さい。

注意：1画像1ページを想定しているためScan Tailorで1画像複数ページを処理している場合は正常に動作しません。

出力のフォルダ構成 (実行後に作成されます)

処理を実行した後は、"cropped"という名称のフォルダを作成し、その下に処理対象フォルダと同じ名前のフォルダと処理したファイルを格納します。

上書き保存するためcroppedという同じ名称のフォルダやファイルがあれば作業前に退避しておいて下さい。



出力されたファイルの幅、高さ、内容を確認して下さい。変な処理をされていた場合は手動で作成しましょう。

補足：0.8以降は、出力フォルダ名を "cropped" から "cropped_<プロジェクトフォルダ名>"に変更しました。(同じcroppedという名称にしていたら、フォルダ移動などでどれがどれだったか見分けがつかなくなったため。)

スクリプトの実行

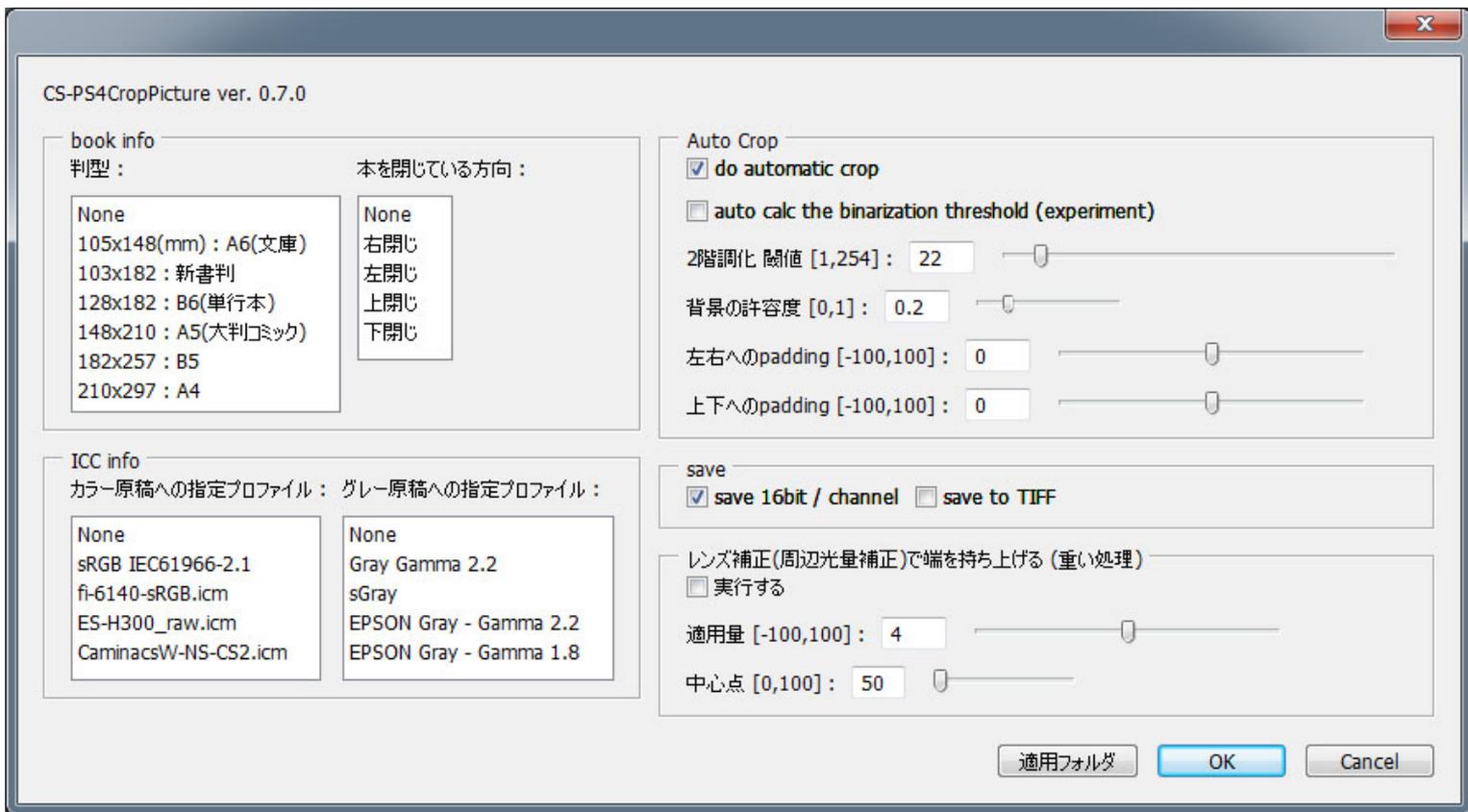
全体的な流れは次のようになります。

1. GUIのスクリプトを起動するとGUIが立ち上がります。
2. GUIからプロジェクトフォルダごとにパラメータを指定した後
3. バッチ処理の実行を開始します。

GUI

GUIのスクリプトを起動するとパラメータを設定するダイアログが表示されます。各種パラメータを指定したら、そのパラメータを適用するプロジェクトフォルダを「適用フォルダ」ボタンを押して指定して下さい。指定したプロジェクトフォルダにパラメータをセーブします。

プロジェクトフォルダを全て指定し終わったらOKを、途中でやめる場合はCancelボタンを押してください。OKの場合、続けてバッチ処理用のダイアログが表示されます。



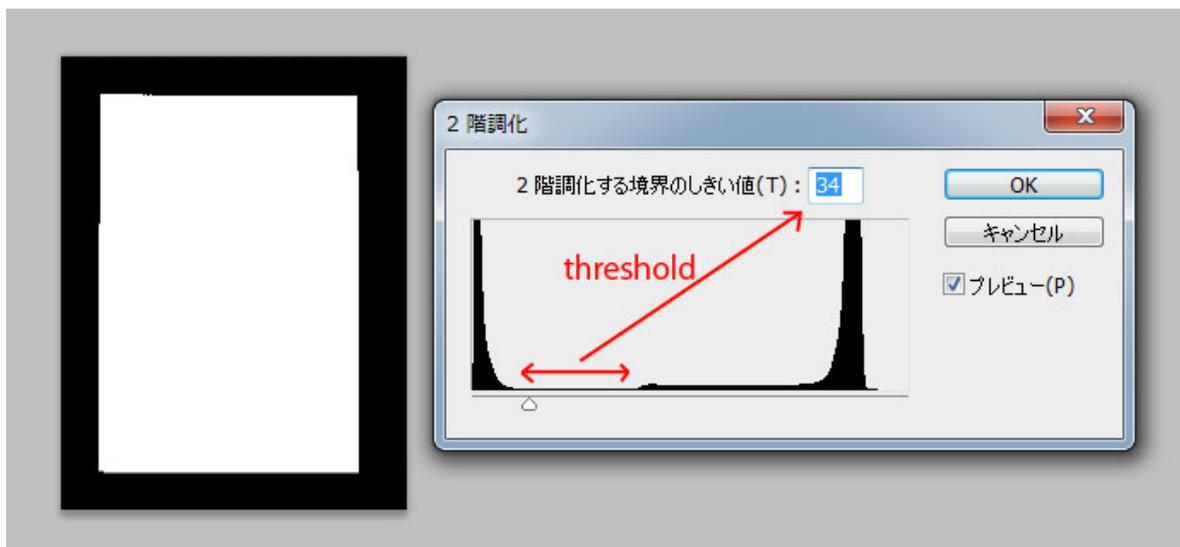
パラメータに関して

大項目	項目	意味
Book Info	判型	原稿サイズの目安。 Auto Cropでの位置決めや自動で閾値を計算させる時にヒントになります。
	閉じている方向	現状使ってません。今後使う可能性があるため残しています。
ICC Info	カラー原稿への指定プロファイル	カラー画像に対してICCプロファイルを指定する場合のプロファイル名。ただしNoneは未指定。 自分で置いたICCプロファイルをスクリプトへ追加して下さい。(下3つは通常的环境にはありません。書き換えの際に利用して下さい。) 存在しないプロファイルを指定しないようにして下さい。
	グレー原稿への指定プロファイル	グレースケール画像に対してICCプロファイルを指定する場合のICCプロファイル名。ただしNoneは未指定。 自分で置いたICCプロファイルをスクリプトへ追加して下さい。(下3つは通常的环境にはありません。書き換えの際に利用して下さい。) 存在しないプロファイルを指定しないようにして下さい。
	do automatic crop	背景除去を行う場合にチェック。
	auto calc the binarization threshold	背景と原稿の見極めに2階調化を使っている。その閾値を自動で算出する場合にチェック。 まだ実験段階で精度がありません。(非推奨)

Auto Crop	2階調化閾値	閾値を自動で算出しない場合に指定する値。2階調化で背景と原稿が分離する点を指定して下さい。
	背景の許容度	背景を残す度合。 0に近づくほど背景部分がなくなります。
	左右へのpadding	検出した左右の辺に対し調整を行います。+は外側へ拡大、内側へ縮小。単位はpixel。
	上下へのpadding	検出した上下の辺に対し調整を行います。+は外側へ拡大、内側へ縮小。単位はpixel。
save	save 16bit /channel	保存画像を各色16bitで保存する場合にチェック。 ただしBMP画像は16bit保存ができません。
	save to TIFF	元のフォーマットのままでなくTIFFフォーマットで保存する場合チェック。
レンズ 補正	実行する	レンズ補正を行う場合チェック。特殊なケース以外は実行をお勧めしない。 紙端が暗くスキャンされる場合に（周辺光量補正で）気持ちだけ持ち上げるために用意しています。
	適用量	周辺光量補正の適用量の値
	中心点	周辺光量補正の中心の値

2階調化の閾値は背景と原稿が分離できる値を指定して下さい。枠がしっかり浮かび上がり原稿内容が飛んだあたりが理想です。

読取りでコントラストが高すぎない場合は、背景の黒部分の山とそれ以外の山になっているはずなのでその間を閾値に指定します。下の例は小説だったのでかなりうまく間が空いていますが、原稿やスキャナによって山が重なりあい、分離がうまくできないものがあります。



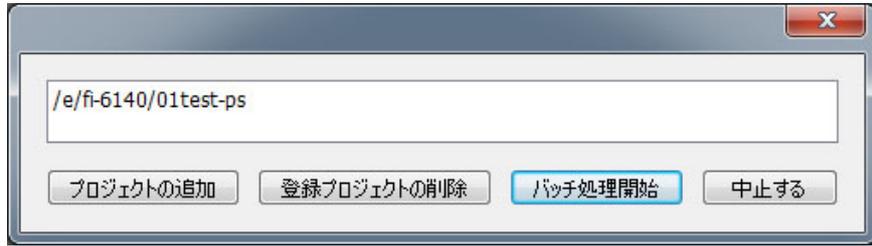
2階調化の閾値は128以上にすると、背景を白と思うように切り替えます。背景の説明で黒と書いてある部分は白と読み替えて下さい。（ただテスト不十分です）

バッチ処理ダイアログ

指定したフォルダで問題なければ「バッチ処理開始」ボタンでバッチ処理を開始します。処理するつもりのないフォルダを登録していた場合は、そのフォルダを選択してから「登録プロジェクトの削除」を押し、バッチ対象から外してください。

パス記法はExtendScriptそのままを使用していますので、表示内容がOSの記法と異なっている場合があります。

下の例ではWindowsの E:\fi-6140\01test-ps というフォルダを意味しています。



別途、適用フォルダでパラメータを保存したプロジェクトフォルダがある場合は「プロジェクトの追加」でバッチ対象に追加できます。バージョンが異なる場合のパラメータ情報の互換性は保証していませんので、同一のバージョンでのみ利用できます。

謝辞

数々の先人の功績と貢献・奉仕に感謝いたします。

Adobe Systems様、Photoshopへスクリプト機能を付与して頂き感謝します。

rajbot様、[Autocrop algorithm for scanned books](#) での成果の公開ありがとうございます。アルゴリズムを勉強させて頂きました。(といっても未熟者ゆえ完璧に理解できず中途半端な理解ですが orz)

古籟一浩様、ASCIIの[古籟一浩のJavaScriptラボ](#)での数々のテクニックの公開、ありがとうございます。実装に行き詰まっていた時に拝見し、このスクリプトを作ることができました。

[Scan Tailor](#)、Joseph Artsimovich様、コントリビュータ各位、
[eTilTran](#)、No.722様、

傾きの自動検出をありがたく使わせて頂いております。傾き検出は一度トライしてみたのですが、最終的にもちはもちやに任せるようにしました。

【ADF】スキャナで連続取り込み017.jpg【OCR】の762様、名も知らぬ者へのご相談と数々のご協力を深く感謝いたします。重ねてお礼申し上げます。

履歴

ver. 0.8

1. 出力フォルダ名を cropped から cropped_プロジェクトフォルダ名 へ変更。
2. 判型にいくつかのサイズを追加。
3. ログファイルが適切に出力されない場合があり修正。

4. ログファイルの記法を見直し。記述位置を改善。幅や高さの差が大きいページを見分けやすいように改善。

ver.0.7

1. 特定の拡張子の場合同じ拡張子で保存するように変更(TIFF, PNG, BMP)
メモ：BMPは24bit保存になるため16bit/channelでは保存できない。PNG保存がなぜか死ぬほど遅い。
2. 他フォーマットをTIFFで保存するオプションを追加。
3. 背景サイズが大きいときにratioに0.2加算でやるとスルーするため調整を行う
4. padding機能(検出位置のオフセット指定)を追加
5. GUIが縦長になりすぎたのでレイアウトを変更。

ver.0.6.1

1. TIFF以外の拡張子がうまくとれてなかったので緊急修正 (thanks 762氏)

ver.0.6

1. 白背景対応 (閾値を128より大きくした場合、背景を白とみるように変更)

ver.0.5

1. GUIを簡易的に国際化対応させる
2. 黒許容度の指定を可能にした。原稿が黒背景だとセンシティブにしたい場合があるため指定可能なように変更
3. スクリプトの名称を変更

ver.0.4

1. 傾き補正支援のためeTilTranのプロジェクト対応

ver.0.3

1. よりバッチ処理させるためGUIと処理を分離。(プロジェクトファイル形式を採用)

ver.0.2

1. 傾き補正に対応。
Scan Tailor (0.9.10)のプロジェクトファイルの読み込みに対応

ver.0.1

1. ボーダー(黒背景)を自動的に取り除く処理の実装

連絡先

もし本プログラムに関して連絡事項がある場合は、こちらの図の緑の円に記述された文字列の2つの★を
"."(ドット)に変更したアドレスへメールして下さい。

内容がプログラムの動作にかかわるもの場合は、以下を明記するようにして下さい。



1. 使用環境：OS、Photoshopのバージョン
2. 使用スクリプトとバージョン
3. 動作内容・現象
4. 実施・再現手順

ちらしのうら (与太話)

多分このツールは普通の人には必要ないものです。これまでスキャナの前稿サイズ自動検知でうまく捌けていた人は本当に用がないでしょうし、足を踏み入れない方が幸せです。いやほんとに。

まだ続きを読んでもという事は既に面倒くさい世界に踏み込んでしまっている方でしょうか。であれば多少は分かってもらえるかもしれない。

自分の場合書籍を電子化してモニターでみるようになったら、傾きが気になってきました。気になったら終わりだと思っても、後引いて、なんとかできんのかのう？と試行錯誤と失敗を重ね、そして処理できない間にずんずん積みあがっていくという、まあどうしようもなくさえない話がバックグラウンドにあります。

スキャナの斜め補正を有効にするととんでもない結果になることが偶にあり、後からその原稿を抜き出してスキャンし直すのも面倒なんで、斜め補正を無効にしました。読み直しを繰り返すと原稿も痛みますし。最近を読み取った後から斜め補正をオート、セミオートでやってくれるScan TailorとかeTilTranといった便利なソフトウェアがあり、助けられております。便利なので是非使ってみてください。

で、斜めってるのを元に戻す際にスキャナがカットしてしまった領域が欲しくなる場合があります。というわけでオーバースキャン設定で読み取るようになりましたが、多めに読み取っているんで今度は領域をクロップしなければなりません。Scan TailorやeTilTranにも保存領域を指定する機能があるのですがいまいち肌にあいませんでした。

[Scan Tailor](#)というのは海外で知名度があるpost-processingツールです。後ろ方のステージで領域指定の部分ができます。

このソフトウェアは原稿をあるサイズにしていくにはいいのですが、雑誌やコミックで原稿をできるだけとりたい場合色々面倒でした。

後、ぴったりさせた原稿の角度補正の動きに不満があります。角度補正した際に足りない領域が補われるんですが、それがなぜかシアンで補完されます。何も縁をそんな目立つ色にせんでもええやん。これだとオーバースキャン、つまりスキャンした原稿の周囲が余分に覆われていれば問題ないのですが、そうでない場合に頭を抱えます。背景色をカスタマイズできれば問題ないのですが。元々がデジカメでスキャンするのを想定しているようでこの辺りの不十分さはしょうがないのかもしれませんが。Forumの過去ログにImageMagickで色変換できるよと返していた人がいたけど……そうじゃねえ、原稿でシアン使ってたらどうすんだ？こいつら分

かってへん、分かってへんがな、と独りたそがれました。いやその質問を投げた人とは一緒に飲める気がするw
角度補正がはまった時はかなりいいので、GUIの使いづらさなど細かい部分でほんと惜しいソフトなんすよねえ……

日本で有名なeTilTranですが、領域の指定がいまいち直観的でなくて、挙動がいまいち分かりづらかったです。私の使い方が悪い
のでしょうけど、数値を入れて間違ったらもう一度入れてっていうのがどうも肌に合いませんでした。GUIで領域を指定できたら
なあとアップデート待ってたんですが、機能アップしそうにないですねえ…

他に吐きだしたもののdpiがどうもスキャンしたものと異なってるようなんですが、大丈夫なんかなあと。メタデータが足りないの
ですかねえ？

そんなわけで出来るだけ原稿を残して、かつ楽なクロップ方法を模索します。

クロップ（日本ではトリミングの方が一般的な言葉なんですか?）って、Photoshopがあれば簡単にできるんちゃうん? だって
プロが常用してるソフトなんですよ、とっていました。今ではそんな甘い昔の自分に蹴り入れたい気分ですがw

勿論、トリミング機能はあるにはあります。が、カットする部分が均一色になっていることが前提になっています。スキャンした原
稿の画像をよくよくみると分かりますが背景は同一色にはならず、トリミングコマンドは自分が思ったように背景を除去してく
れませんでした。

Photoshopには他にも「角度補正して切抜き」という、背景から写真を切り抜いてくれる機能があるんですが、これを書籍をスキャン
した画像に実行したらもの見事に幅や高さがばらばらになりました。当初は幅や高さが違いすぎるのをはねてしまう事を検討
したのですが、最終的にあまりにも不揃い過ぎてあてにならんと感じました。正直、全く調整する部分がないというのもどう
かと思う……

それでも諦めきれずに試行錯誤を行いまして、

ADFでスキャンしたのならほぼ同一の位置だろうと領域を指定してクロップするスクリプトを組んでみましたが、まあやっぱり搬送
ずれがあって気になりだすと気になってくる。それよりも領域を指定するのが一々めんどい。

色域指定でうまく指定できればと思ったけど、なかなか背景との分離がうまくできない。

Paul Jaruszewskiさんが[AutoCrop Script](#)を公開してくれてますが、これは対象がスキャン画像でなくて、ツールなどで傾けた際に
補う均一色が前提に見えるなあ…

等々、結局うまくいきませんでした。Photoshop以外にも対象を広げても芳しくありません。

Scan Tailor、eTilTranは前出の通り。他は、

XnViewでオートクロップが行えますが、安全マージンに倒すみたいで安心はできるんだけど結構縁に残った黒が目立つ……

ImageMagickは、スクリプトがWindowsでうまく動作しなかったり、自分で組む価値あるのか調べるために動作を手動で拾って
いったら中間ファイルのやりとりでとても遅そうなのが分かったり…

E.C. う〜ん手間かかるなあ。オートにならんか。どのみちチェックしなきゃならんのなら最初からというのも分けるけどかなり負担
かかるなあ…

……と、まあ作者の数年来の無駄な試行錯誤の果てにこのスクリプトにたどり着いています。他のアプリだとどうもなあという場
合に使用してみてください。

しかしauto crop方法をたまたま検索してたのですが、質問が多い割にはこれだという解がないのがどうにもこうにも…… AdobeさんSなの? それとも私が無知なだけで常識的なTipsがあるんでしょうか。

そういえばこの前Adobe製品のアップグレードポリシーの変更がありました。正直どうしたもんかと。資金的に毎回バージョンアップさせるのがつらいっすよ…… > Adobe様

前にいいおっさんが銀行残高6桁切った事があるほどの体たらくなので無い袖は振れませぬ。しかしあの時はつらかったなあwww orz

まじめな話全てのバージョン、OS含む環境でテストなんて資金的にも時間的にも無理なので、使えればラッキー、割切れるんで構わんと思う人だけが使って下さい。その辺の塩梅はお願い致します。

2011/12/25

つぶ@coffee shop 拝

(2012/05/01誤字を修正)