PDF3BAR PDFデータからバーコード情報読み取り

Kernel Computer System
カーネルコンピュータシステム株式会社

本社：パッケージ販売部
〒221-0056
横浜市神奈川区金港町 6-3 横浜金港町ビル
TEL：045-442-0500　FAX：045-442-0501
URL：https://www.kernelcomputer.co.jp

PDF データ内に存在するバーコードから情報を読み取り、CSV ファイルで出力します。
また、バーコードを含む画像を抽出することも可能です。

<PDF3BAR の概要図>

動作環境

対応 OS：
弊社製の PDF から画像変換プログラムや PDF 編集用 DLL を使用時

Windows 10
Windows 11
Windows Server 2016
Windows Server 2019
Windows Server 2022
Windows Server 2025

任意のコマンドラインアプリケーションを使用時は、上記の環境に加えて使用するアプリケーションの対応 OS に依存します。

標準価格

PDF3BAR　　：　35 万円/本 （税抜き）

特徴

・認識可能なバーコード。
CODE39、CODE128、GS1-128(UCC/EAN-128)、JAN/EAN-13、JAN/EAN-8、ITF、NW-7、
郵便カスタマバーコード、QR コード、マイクロ QR コード、PDF417、GS1 DataBar(RSS)、
GS1 合成シンボル(EAN.UCC 合成シンボル)、DataMatrix(ECC200 のみ)
・読み取り範囲(全景も可能)指定が可能。
・コマンドラインからバッチ変換が可能。同一フォルダなら一括して全て変換可能。
・属性ファイルでバーコードを認識する場合にさまざまな設定が可能です。
①検出するバーコードの桁数を指定。
②検出するバーコードの桁数の範囲を指定。
③検出するバーコードの各桁の文字種を指定。
④出力するバーコード情報の書式を指定。
⑤スタート/ストップキャラクタの取り扱いを指定。
⑥ディザノイズの除去。
・バーコード検出順位方向の指定に対応。
・分割 QR コードの連結。
ページごとのデータの座標から、バーコードの順番をソート(上下左右)。
・PDF 内の画像から、バーコードを含んだものを TIFF 形式で抽出。
・標準のパスワードによるセキュリティであれば解除して実行が可能。
※パスワードが事前に分かっている場合のみ
・自動正立
・全ページ対象(最初または直前のバーコードの向きで、ページの回転角度を決定)
・バーコードのあるページが対象(バーコードがあれば、その向きから回転角度を決定)
・PDF データのページ削除(全ページの削除は不可)。 ※2
・バーコードを含む、もしくは、含まないページを削除。
・白紙ページ(バーコード情報が BLANK)の削除 ※1
・バーコード読み取り結果によって、PDF データのページを分割 ※2
・バーコードが検出されたページごと
バーコードが検出されたページを先頭ページとして、ページを分割。
・バーコードの内容ごと
ページ順に解析し、バーコードの内容が変わったページを先頭ページにしてページを分割。
・バーコードの内容ごと(順不同)
バーコードの内容ごとにページをまとめて、分割した PDF を出力。
・バーコード読み取り結果を使用して、入力 PDF データのファイル名を変更が可能。
・Adobe Acrobat は不要です。
・GhostScript などのコマンドラインアプリケーションを使用して PDF を読み込むことが可能。
※事前にインストールおよび設定をしておく必要があります。
※入出力ファイル名が指定可能で、かつ、マルチページのラスターに出力できる必要があります。

※1. 実行する条件によっては対応不可。
※2. ラスター化を行う方法によって制限が異なります。

操作方法

PDF3BAR は、コマンドラインアプリケーションです。通常のコマンドと同じように、コマンドラインから実行し、バッチファイルに組み込むことができます。
「コマンドプロンプト」で DOS ウィンドウを開いて、その中で実行します。

属性ファイルの設定

変換を始める前に、変換に必要な情報を属性ファイル上で設定します。
以下の属性ファイルでバーコード認識をする際に必要な情報を設定します。
通常、基本的な初期値があらかじめ設定されています。
pdf_bar_in.atr ・・・・・・・・・・・ PDF 入力設定属性ファイル
pdf_bar.atr ・・・・・・・・・・・・・ バーコード認識設定属性ファイル

実行例

バーコード読み取りを実行するには、次のような形式でコマンドラインに入力します。

```
実行モジュール名　入力ファイル -O 出力ファイル(*.csv)　-各オプション
```

例１：　入力ファイル[BarSample.pdf]の全範囲を検索し、通常の方法で全てのタイプのバーコードを読み取り、バーコード情報を出力ファイル[tes.csv]に出力します。

```
>PDF3BAR　BarSample.pdf -Otes.csv
```

PDF3BAR　　：実行モジュール名
BarSample.pdf　：入力 PDF データ名
tes.csv　　：バーコード情報を出力するファイル名

-O　　：変換結果を出力するファイル名を指定するオプション。
マイナス記号とアルファベットの「オー」です。
-O と変換後のファイル名の間には空白を入れないようにします。
変換前のファイル名と変換後のファイル名に同じものを指定することはできません。

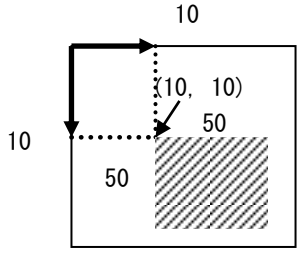
※オプション指定は大文字でも小文字でもかまいませんが、全角文字は使えません。
※オプションの詳細についてはオプションの項目を参照ください。

例２：　入力ファイル[BarSample2.pdf]の指定された読み取り検索領域を検索し、縦方向に Code128 のバーコードの読み取りを行い、バーコード情報を出力ファイル[tes2.csv]に出力します。

```
>PDF3BAR BarSample2.pdf -Otes2.csv -BDIR90 -BTYPcode128 -BSX10 -BSY10 -BSW50 -BSH50
```

PDF3BAR　　：実行モジュール名
BarSample2.pdf　：入力 PDF データ名
tes2.csv　　：バーコード情報を出力するファイル名

-O　　：変換結果を出力するファイル名を指定するオプション。
-BDIR90　　：縦方向バーコードを読み取ります。
-BTYPcode128　：Code128 バーコードを読み取るよう指定します。
-BSX10 -BSY10 -BSW50 -BSH50　：左上から右に 10mm、下に 10mm の地点を起点にし、
幅 50mm、縦 50mmの範囲にあるバーコードを検索します。



例３：　入力ファイル[BarSample3.pdf]からバーコード情報を持つ画像を抽出します。
※抽出された画像は TIFF 形式となります。

```
>PDF3BAR　BarSample3.pdf　-BOtes.tif
```

PDF3BAR　　：実行モジュール名
BarSample3.pdf　：入力 PDF データ名

-BO　　：画像を出力するファイル名を指定するオプション。

実行時のオプション

PDF3BAR には、変換時に指定できるさまざまなオプションが用意されています。
・オプションは、「-」で始まります。
・オプションには、大文字/小文字の区別はありません。
・パラメータがある場合は、空白を開けずにオプションの直後に指定します。
（例）オプションが「-O」、パラメータが「file」の場合：
（正）-Ofile
（誤）-O file
・複数のオプションを同時に指定する時は、それらを空白で区切ります。
・オプションを指定する順番は、処理の内容には影響しません。

入出力ファイルに関するオプション

-O[filename] ： 解析したバーコード情報を出力するファイル名を指定します。
※デフォルトの出力先は標準出力です。

-BO[filename] ： バーコードを含む画像抽出用の出力ファイル名を指定します。
このファイル名にページ番号やインデックスが付加されて出力されます。
（例：outdat_ページ_1_画像_0006.tif）

-M ： このオプションを指定しておくと、-O で指定した出力ファイルが既に存在する場合、バーコード情報はそのファイルの最後に追加されます。
デフォルトでは、上書きされます。

-U[in_atr_name] ： 入力属性ファイル名（初期値：pdf_bar_in.atr）
入力ファイルのデコード処理用のパラメータを設定するファイル。

-Z[atr_name] ： 出力属性ファイル名（初期値：pdf_bar.atr）
バーコード認識処理用のパラメータを設定するファイル。

-ASSORT[フォルダ名] ： PDF をバーコードの解析結果に応じてページを分割します。
その際に分割後の PDF データを出力したいフォルダ名を指定します。
出力ファイル名は、最初に読み取られたバーコードの内容を使用します。
バーコードが検出されなかった場合のファイル名は、「NODATA.pdf」となります。
分割の設定は、属性ファイルで行います。

バーコード読み取りに関するオプション

-BDIR[dir] ： 読み取るバーコードの方向を指定します。

-BDIR1 または -BDIR0
（初期値）

-BDIR2 または -BDIR180

-BDIR4
または
-BDIR270

-BDIR8
または
-BDIR90

- 指定できる方向は、90 度単位の 4 方向（0 度, 90 度, 180 度, 270 度）です。
- 複数の方向を組み合わせて指定することもできます。
（例）-BDIR0,90,180,270　　（4 方向全てを指定）
- 検出するバーコードの方向は、出力属性ファイルの BAR_DIRECTION 属性でも指定できます。スキャン時に発生する傾き程度なら、特に影響はありません。
- 1 次元バーコードについては、縦方向（90 度、270 度）よりも横方向（0 度、180 度）の検出の方が高速です。
- 2 次元コードの場合（GS1 合成シンボルを除く）は、この指定にかかわらず、全ての方向で検出されます。

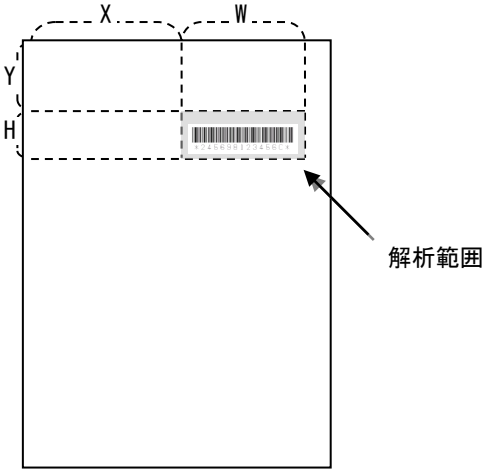
-BTYP[typ] ： 読み取るバーコードのタイプを指定します。

-BTYP**JAN13**	: JAN/EAN 13 桁
-BTYP**JAN8**	: JAN/EAN 8 桁
-BTYP**JAN**	: JAN/EAN 13 桁または 8 桁
-BTYP**CODE39**	: Code39
-BTYP**CODE128**	: Code128
-BTYP**ITF**	: ITF
-BTYP**NW7**	: NW-7
-BTYP**EAN128**	: GS1-128(UCC/EAN-128)
-BTYP**CUSTOMER**	: 郵便カスタマバーコード
-BTYP**RSS**	: GS1 DataBar(RSS)
-BTYP**QR**	: QR コード
-BTYP**MQR**	: マイクロ QR コード
-BTYP**PDF417**	: PDF417
-BTYP**DM**	: DataMatrix(ECC200)
-BTYP**CS**	: GS1 合成シンボル(EAN.UCC 合成シンボル)
-BTYP**BLANK**	: 白紙ページの検出
-BTYP**ANY**	: GS1 DataBar(RSS)、PDF417、CS 以外の全て(初期値)
-BTYP**ALL**	: 全て

複数のタイプを指定することもできます。この場合、名称を「,」で区切って複数指定してください。
(例) JAN と ITF の場合: -BTYPJAN,ITF

-BSX[X], -BSY[Y] ： バーコードの読み取り開始位置を指定します。
左上を起点とし、右に Xmm、下に Ymm の地点を開始位置とします。

-BSW[W], -BSH[H] ： バーコードの読み取り範囲を指定します。
-BSX ,-BSY で指定された地点から幅 Wmm、高さ Hmm の範囲を検索範囲とします。



※ -BSX, -BSY, -BSW, -BSH の指定がない場合、入力イメージの全範囲を検索範囲とします。

-N[page] ： 入力がマルチページファイルの場合、処理するページを指定します。

-N0	: 全ページを処理(初期値)
-N3	: 3 ページ目のみ処理
-N3-5	: 3 ページ目から 5 ページ目まで処理
-N-5	: 1 ページ目から 5 ページ目まで処理
-N5-	: 5 ページ目から最終ページまで処理
-N3,5	: 3 ページ目と 5 ページ目を処理

セキュリティに関するオプション

-UNLOCK[pass_file]　：　パスワード解除用ファイル名を指定します。
テキスト形式のファイル内にセキュリティ解除用のパスワードを記述します。
解除可能なのはパスワードによるセキュリティのみです。

バッチ処理に関するオプション

-BATCH[batch_file]　：　ファイルの一括処理を行う、バッチファイル名を指定します。
指定されたファイルまたはディレクトリ内のファイルを一括処理したい時に利用します。

制限事項

◆ **動作条件について**

PDF3BAR は PDF をラスター変換するアプリケーションの機能を利用して TIFF 形式のイメージを出力し、ラスター内のバーコード解析を行っています。そのため、使用するアプリケーションによって画像化ができない PDF ファイルやページや、オープンできない PDF ファイルにつきましては対応不可となります。

PDF3BAR に付属している PDF 編集用の DLL や画像変換用のプログラムを単独で使用することはできません。

用紙サイズ、ページ数、TIFF 化する時のフォーマットや解像度によっては、中間 TIFF ファイルのサイズが大きくなり、HDD 等の空き容量が不足して正常に解析ができない場合があります。そのような場合は解像度を下げる等を行ってください。

※テンポラリフォルダを使用します。入出力ファイルのあるドライブとは限りません。

Windows の OS を長時間(数ヶ月単位)再起動せずに使用すると環境により、Windows の動作が不安定になる場合があります。これは、弊社アプリケーションのみの問題ではなく、Windows を含む、利用される全てのソフト・ドライバの複合要因であり、発生した場合は原因の特定および回避は全てのアプリケーションが関係しているため、現状では対応が不可能です。長期間連続運用を行う場合、上記の問題が発生することがありますので、定期的な Windows 再起動を奨励致します。

◆ **バーコードの判定について**

TIFF 形式のイメージを出力し、そのイメージに対してバーコードの解析をしています。そのため、用紙サイズが大きい場合や、セキュリティが設定されている場合など、画像化ができない PDF ファイルにつきましては対応不可となります。また、画質や大きさによってはバーコードとして認識できない場合もあります。
PDF の画像化に弊社の PDF 編集用 DLL を指定した場合は、バーコードが画像であれば検証の対象となります。画像以外の図形で作成されているバーコードは認識できません。また、画像の形式によっては対応できない場合があります。

◆ **セキュリティについて**

セキュリティが設定されている PDF ファイルにつきましては、基本的に未対応となります。
標準セキュリティのパスワードによるセキュリティであれば、「-UNLOCK」オプションにより、解除をして処理を行うことは可能です。
ただし、パスワード付きのセキュリティの解除には、パスワードが事前に分かっている必要があります。

◆ **バーコードの認識について**

- キャラクタ構成が規格に準拠していないバーコードは検出できません。
- イメージに含まれるバーコードに似たパターンを、短いバーコードとして検出してしまうことがあります。このような現象の多くは、検出するバーコードの種別や方向を制限する、BAR_DIGITS で桁数の範囲を指定することにより回避できます。
- 1 次元バーコードは少し斜めになっていても認識することが可能なので、用紙スキャン時の傾き程度であれば問題ありませんが、斜め 45 度など大きく傾いている場合は認識できません。
- 各アプリケーションで PDF を正常な TIFF に変換できない場合は解析できません。
- バーコードにはクワイエットゾーン(周囲に設ける空白の領域)が必要です。特に 1 次元バーコードの左右に十分なクワイエットゾーンがない場合は認識することができません。
- 写真などの網点や、図のハッチング部分をバーコードと誤認することがあります。このような誤認識を防ぐには、BAR_INTERVAL に網点などのピッチよりも大きい値を指定してください。
- 弊社の PDF 編集用 DLL を使用する場合は、バーコードが画像であれば検証の対象となります。画像以外の図形で作成されているバーコードは認識できません。また、画像が使用されていないページは白紙として判定されます。

◆ **JAN について**

- チェックデジットの部分もデータとして出力されます。
- EAN コードも、バーコード種別上は JAN と表示されます。
- UPC-A は、先頭に「0」が付加された 13 桁の EAN コードとして認識できます。UPC-E は認識できません。
- 解像度が適切でない場合、間違った内容のデータが検出される場合があります。

◆ **Code128 について**

検出するバーコードのタイプとして EAN128 または GS1-128 が指定されていて、スタートキャラクタの次に FNC1 のある Code128 のバーコードを検出した場合、それを GS1-128(UCC/EAN128)と認識し、以下の処理が行われます。

- スタートキャラクタの次の FNC1 は削除。
- 2 個目以降の FNC1 は GS 文字(ASCII コード 29)に置換。
- FNC2～FNC4 は削除。

上の条件に当てはまらない場合、Code128 に含まれる FNC1～FNC4 は、全て無視されます。

◆ **ITF について**

- ITF_MIN でも最低桁数を指定できます。
- 解像度が適切でない場合、間違った内容のデータが検出される場合があります。

◆ NW-7 について
・NW7_MIN でも最低桁数を指定できます。
・解像度が適切でない場合、間違った内容のデータが検出される場合があります。
・キャラクタ間ギャップが大きすぎると正しく認識されない場合があります。

◆ Code39 について
・解像度が適切でない場合、間違った内容のデータが検出される場合があります。
・キャラクタ間ギャップが大きすぎると正しく認識されない場合があります。

◆ GS1 DataBar(RSS)について
・チェックデジット、アプリケーション識別子の部分もデータとして出力されます。
・イメージに含まれるバーコードに似たパターンが極端に多い場合は、バーコードが正しく認識されない場合があります。

◆ QR コードについて
・明暗反転には対応していません。

◆ マイクロ QR コードについて
・漢字などの複数バイト文字はシフト JIS でエンコードされているものとします。
・明暗反転および表裏反転には対応していません。

◆ PDF417 について
・マクロ PDF417 には対応していません。
・漢字などの複数バイト文字はシフト JIS でエンコードされているものとします。

◆ DataMatrix について
・ECC200 のみ対応しています。
・漢字などの複数バイト文字はシフト JIS でエンコードされているものとします。
・明暗反転および表裏反転には対応していません。
・次のコード語には対応していません。
FNC1(232)、増築(233)、読み込みプログラミング(234)、
05 マクロ(236)、06 マクロ(237)、ECI(241)
・以下の条件に当てはまる場合、正しく検出されない場合があります。
・1 セルの 1 辺のサイズが小さいもの
・90 度単位から大きく傾いているもの
・ゆがみがあるもの
・セルの形状が正方形でない

◆ GS1 合成シンボル(EAN.UCC 合成シンボル)について
・UPC-E をリニアコンポーネントとするシンボルは認識できません。

◆ PDF の読み込みに弊社の PDF 編集用 DLL を使用した場合について
・以下のような PDF に関しては対応できません。
・ 電子署名の付与された PDF
・ PDF ポートフォリオ
・ PDF 2.0 の範囲を超える、アプリケーション独自の拡張が使用された PDF
・ バーコードが画像であれば検証の対象となります。画像以外の図形で作成されているバーコードは認識できません。
・バーコードの画像かどうかにかかわらず、なんらかの画像が存在しないページは白紙として認識されますので、白紙ページの判定や削除に関する機能は対応できません。
・画像であっても、次の圧縮形式が使用されているイメージは認識できません。これらの画像を含むデータが指定された場合はエラーとなります。
・JPXDecode(JPEG 2000)のうち拡張機能が使用されているもの。
・PDF 上で回転・歪みを使用して表示されているバーコード画像は、正常に認識できない場合があります。

◆ PDF の読み込みに弊社の PDF から画像変換のプログラムを使用した場合について
・以下のような PDF に関しては対応できません。
・ 電子署名の付与された PDF
・ PDF ポートフォリオ
・ PDF 2.0 の範囲を超える、アプリケーション独自の拡張が使用された PDF
・ ページの縦か横が 10m を超える PDF
・ 2GB を超えるデータ。
・次の圧縮形式が使用されているイメージは認識できません。これらの画像で作成されたバーコードは認識できません。また、これらの画像を含むページは白紙検出が正常に行えない場合があります。
・JPXDecode(JPEG 2000)のうち拡張機能が使用されているもの。

◆ **PDF の読み込みにコマンドラインアプリケーションを使用した場合について(PDF_MODE=5)**

・使用するアプリケーションで読み込むことができない PDF に関しては、対応できません。
・使用するアプリケーションで画像に変換できない PDF には対応できません。
・複数ページの PDF がマルチページの画像に変換される必要があります。
・出力ファイル名が指定したままの名前にならない場合は対応できません。たとえば、自動的にファイル名に番号などが付加される場合は対応できません。

◆ **PDF の編集を伴う処理について**

・以下のような PDF に関しては対応できません。
  ・ 電子署名の付与された PDF
  ・ PDF ポートフォリオ
  ・ PDF 2.0 の範囲を超える、アプリケーション独自の拡張が使用された PDF
・バーコード画像の抽出につきましては、次の圧縮形式が使用されている場合は抽出できません。
  ・JPXDecode(JPEG 2000)のうち拡張機能が使用されているもの。