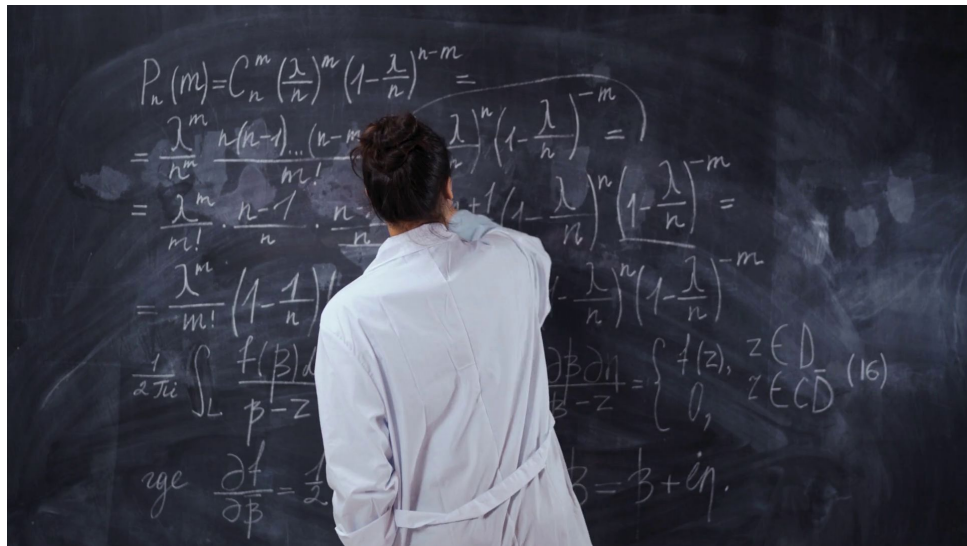


動画の再生情報を付与した PDFの自動生成

奥田 遼

1. 背景



- 授業でわからない部分だけ確認したい



- 動画を撮って、飛ばしながら見る



- 板書との対応がわからず、**手動で再生位置を探す**必要がある
 - 面倒

2. 目的・条件

- 目的

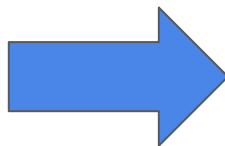
- 授業の板書をクリックしたら、動画の再生位置を変更する

- 条件

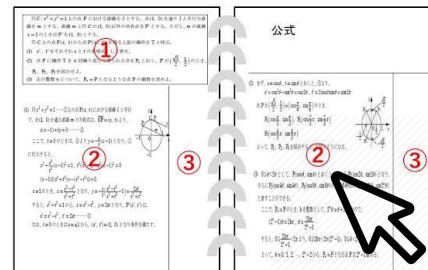
- 板書は見やすくするためにPDF形式等でまとめる

3. 入出力

入力
動画



出力
PDF + 再生機能



4. システム概要

入力

ソフトウェア

出力

1. 読み込み

2. 文字抽出

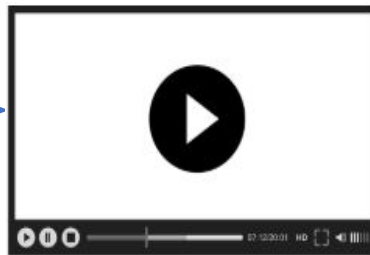
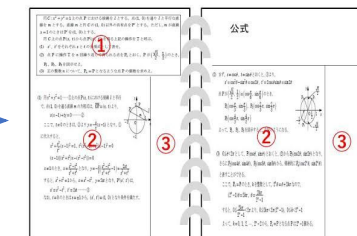


4. 再生位置変更



5. 出力

3. PDF生成



5. 使用技術 (1/3)

OCR

出典: フリー百科事典『ウィキペディア (Wikipedia)』

OCR

- **光学文字認識** (Optical Character Recognition) の略称。
 - **光学式文字読取装置** (Optical Character Reader) の略称。
- **オイル循環率** (Oil Circulation Ratio) の略称。
- **過電流継電器** (Over Current Relay) の略称。



このページは**曖昧さ回避のためのページ**です。一つの語句が複数の意味・職能を有する場合の水先案内のために、異なる用法を一覧にしてあります。お探しの用語に一番近い記事を選んで下さい。このページへリンクしているページを見つけたら、リンクを適切な項目に張り替えて下さい。

カテゴリ: 曖昧さ回避

● 文字抽出

- Tesseract + PyOCRを使用
- 画像から文字を抽出する
- 改行などの**構造を保持**
- 以下はwikipediaでの実行例

出典: フリー百科事典『ウィキペディア (Wikipedia)』

OCR

- ・ **光学文字認識** (Optical Character Recognition) の略称。
- ・ **光学式文字読取装置** (Optical Character Reader) の略称。
- ・ **オイル循環率** (Oil Circulation Ratio) の略称。
- ・ **過電流継電器** (Over Current Relay) の略称。

このページは**曖昧さ回避のためのページ**です。一つの語句が複数の意味・職能を有する場合の水先案内のために、異なる用法を一覧にしてあります。お探しの用語に一番近い記事を選んで下さい。このページへリンクしているページを見つけたら、リンクを適切な項目に張り替えて下さい。

カテゴリ: 曖昧さ回避

5. 使用技術 (2/3)

円 $C: x^2 + y^2 = 1$ 上の点 P における接線を l とする。点 $(1, 0)$ を通り l と平行な直線を m とする。直線 m と円 C の $(1, 0)$ 以外の共有点を Q とする。ただし、 m が直線 $x = 1$ のときは Q を $(1, 0)$ とする。

円 C 上の点 $P(x, y)$ から点 Q までの距離を上記の操作を n 回繰り返すと、

(1) x', y' をそれぞれ x と y の n 乗とする。

(2) 点 P に操作 T を n 回繰り返して得られる点を Q_n とおく。 P が $(\frac{\sqrt{3}}{2}, \frac{1}{2})$ のとき、 Q_1, Q_2 を求めよ。

(3) 正の整数 n について、 $Q_n = P$ となるような点 P の個数を求めよ。

① 円 $x^2 + y^2 = 1$ 上の点 $P(x, y)$ における接線 l と平行で、点 $(1, 0)$ を通り直線 m の方程式は、 $\overrightarrow{OP} \cdot \vec{m} = (x, y) \cdot \vec{m}$ となり、 $\vec{m} \cdot (x-1, y) = 0$ である。

ここで、 $(x-1) + y = 0$ となる。

に代入すると、 $x^2 + y^2 - (x-1) + y = 0$ となり、 $x^2 + y^2 - x + y + 1 = 0$ となる。

$x=1$ のとき、 $y^2 = 0$ となり、 $y=0$ となる。 $x=1, y=0$ のとき、 $x^2 + y^2 = 1$ となる。

したがって、 $x^2 + y^2 = 1$ 上の点 $P(x, y)$ に対して、 $Q(x', y')$ は、 $x' = x^2 - y^2, y' = 2xy$ となる。

また、 $(x-1) + y = 0$ となる点 $P(x, y)$ に対して、 $Q(x', y')$ は、 $x' = x^2 - y^2, y' = 2xy$ となる。

② P が $(\frac{\sqrt{3}}{2}, \frac{1}{2})$ のとき、 Q_1 を求めよ。

$x' = x^2 - y^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}, y' = 2xy = 2 \cdot \frac{\sqrt{3}}{2} \cdot \frac{1}{2} = \frac{\sqrt{3}}{2}$ となる。

点 Q_1 は $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ となる。

Q_2 を求めよ。

$x'' = x'^2 - y'^2 = \frac{1}{4} - \frac{3}{4} = -\frac{1}{2}, y'' = 2x'y' = 2 \cdot \frac{1}{2} \cdot \frac{\sqrt{3}}{2} = \frac{\sqrt{3}}{2}$ となる。

点 Q_2 は $(-\frac{1}{2}, \frac{\sqrt{3}}{2})$ となる。

③ 正の整数 n について、 $Q_n = P$ となるような点 P の個数を求めよ。

① $Q_1 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

したがって、 $Q_1 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

② $Q_2 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

したがって、 $Q_2 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

③ 正の整数 n について、 $Q_n = P$ となるような点 P の個数を求めよ。

① $Q_1 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

したがって、 $Q_1 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

② $Q_2 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

したがって、 $Q_2 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

③ 正の整数 n について、 $Q_n = P$ となるような点 P の個数を求めよ。

① $Q_1 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

したがって、 $Q_1 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

② $Q_2 = P$ となる点 P を求めよ。

$x = x^2 - y^2, y = 2xy$ となる。

$y = 2xy$ より、 $y(2x - 1) = 0$ となる。

$y = 0$ のとき、 $x = 1$ となる。

$2x - 1 = 0$ のとき、 $x = \frac{1}{2}$ となる。

$x = \frac{1}{2}$ のとき、 $y^2 = 1 - x^2 = \frac{3}{4}$ となる。

$y = \pm \frac{\sqrt{3}}{2}$ となる。

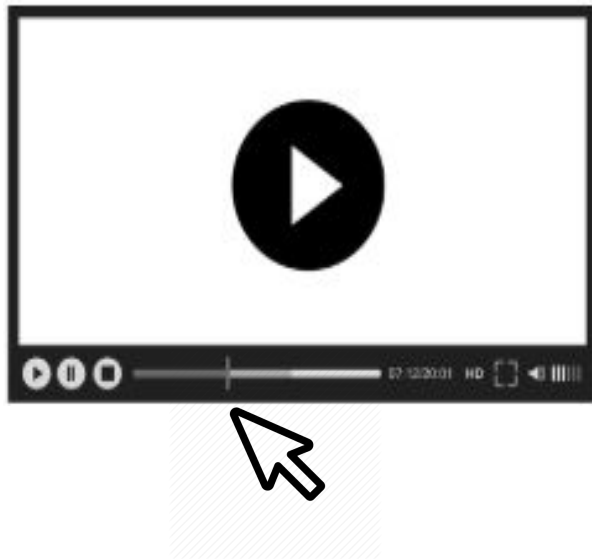
したがって、 $Q_2 = P$ となる点 P は $(1, 0), (\frac{1}{2}, \frac{\sqrt{3}}{2}), (\frac{1}{2}, -\frac{\sqrt{3}}{2})$ の3点である。

③ 正の整数 n について、 $Q_n = P$ となるような点 P の個数を求めよ。

● PDF生成

- ReportLabを使用
- テキストデータを元にPDFを作成する
- 一定の**クラスタごと**に再生位置の情報を付与

5. 使用技術 (3/3)



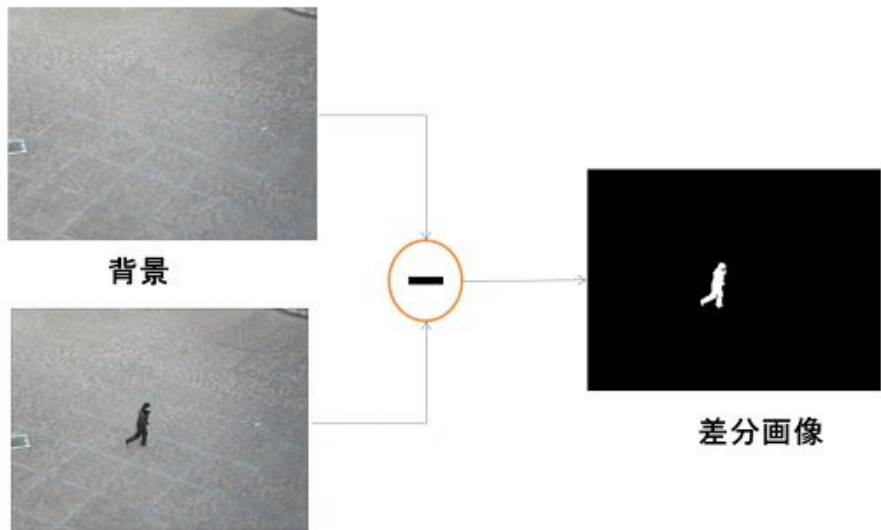
- **再生位置変更**
 - pygameletを使用
 - mp4を音声つきで再生
 - OpenCVでは画像のみ
 - PDFに付与された時間へと再生位置を変更する

6. ディスカッション (アプローチの検討 1/2)



- 現在のアプローチ
 - 文字抽出の結果を元に出力を行う
- **メリット**
 - 改行などを自動で抽出してくれるので、**クラスタ分割が容易**
 - 文字情報が手に入るのでPDF作成が**容易**
- **デメリット**
 - 精度が出ない可能性がある
 - ライブラリを利用しているので **精度改善が難しい**

6. ディスカッション (アプローチの検討 2/2)



- 別のアプローチ (候補)
 - 個別に文字を検出せずに、フレーム間の差分などから板書情報を取得
- メリット
 - 自ら実装できる部分が多く **チューニングが容易**
- デメリット
 - 画像情報のみから **適切なクラスタに分割するのが難しい**
 - 文字情報が取得できず、 **PDF生成が難しい**

6. ディスカッション (出力の検討)

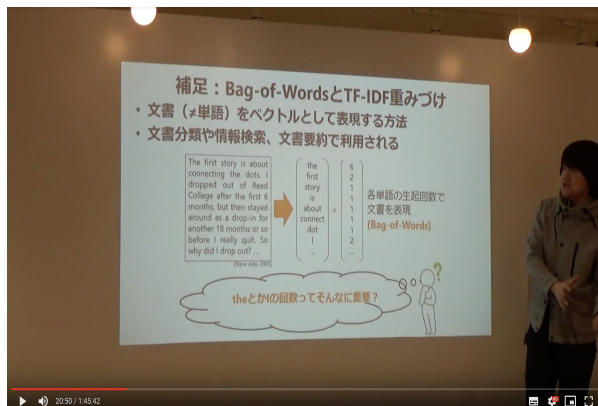
補足: Bag-of-WordsとTF-IDF重みづけ

- 文書 (≠ 単語) をベクトルとして表現する方法
- 文書分類や情報検索、文書要約で利用される

- 綺麗なPDFを出力するのが難しかった場合
- 最優先の機能は動画の再生位置を自動で変えること
 - PDFでなくても良いはず
- 出力方式を変える

例: スクリーンショット

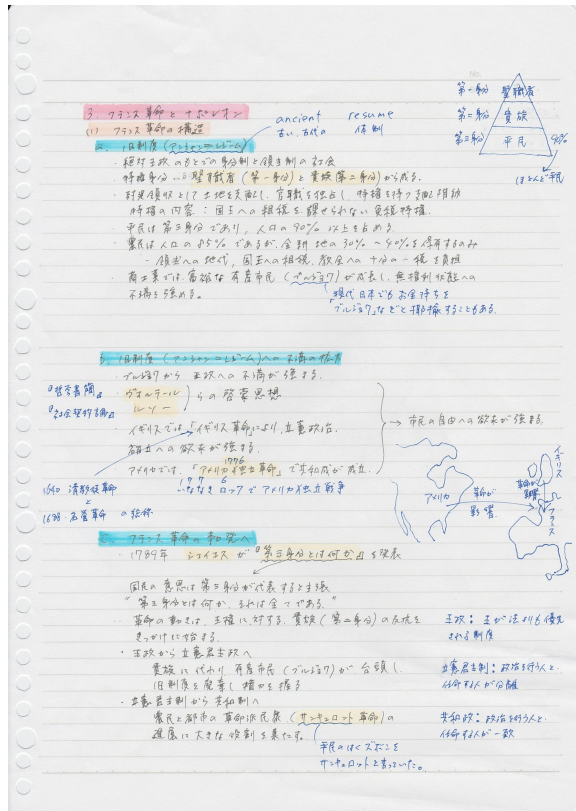
- 一定の条件 (板書の1/4が埋まった場合など) 下でスクリーンショットを取得
- 各スクリーンショットに時間情報を付与し、出力



6. ディスカッション (システムの検討)

以下のような代替案も考えられる

- **アイデア: 入力にノートを加える**
 - ノートから検出した文字・数式を動画内から見つけ、再生位置の情報を付与する
 - PDFとして出力
- **メリット**
 - 時系列を有する探索問題に帰着する。アルゴリズムさえ組めればいいので **解きやすい**
- **デメリット**
 - ノートを用意する **手間がかかる**
 - 大学生はノートを取らないことが多い



7. スケジュール

項目は以下の通り



～ 2 / 6

- pyglet などライブラリを動かしてみる
- テスト用の動画を決定

● ～ 2 / 13

- ディスカッションで示した案のうち、どれを採用するか決定

● ～ 2 / 20

- 適当な再生位置情報をPDFに付与

● ～ 3 / 12

- 動画から再生位置を取得
- 発表準備

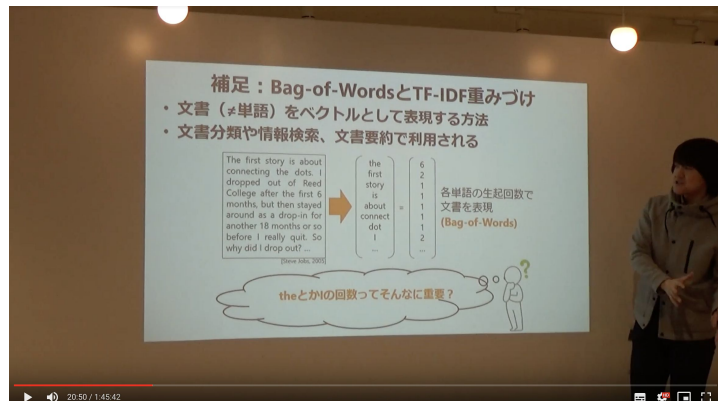
● ～ 3 / 19

- +α

7. まとめ

1. 背景
 - a. 授業の必要な部分だけ簡単に確認したい
2. 目的・条件
 - a. PDFをクリックしたら、動画の再生位置を変更する
3. 入出力
 - a. 動画 → PDF + 再生位置変更機能
4. システム概要
 - a. 入力から出力までの流れについて紹介
5. 使用技術
 - a. 各コンポーネントで用いる技術について紹介
6. ディスカッション
 - a. アプローチについて検討
 - b. 出力について検討
 - c. システムについて検討

Appendix.A PyOCRの実行結果



Bag-of-Words と TF-IDF 重みづけ、

= け

・ 文書 (x え 単語) をベクトルとして表現する方法
` 文書分類や情報検索、文書要約で利用される

- 照明が薄暗い状態でもある程度検出できている
 - ≠などは認識を間違えている
- 実際には右上の出力以降エラー文が大量に続いており、かなりの **エラー処理が必要**

Appendix.B PDFへの情報付与

- pynputを使えばmouseについて監視が可能
- 「python monitor pdf click event」などの直接的な検索では ✕
- 最も妥当そうな実装は以下の通り
 - PDFをWebブラウザ上で表示
 - JQueryでクリックされた座標を取得
 - 事前に定義されたクラスタ領域のいずれに属するか判定
 - JQuery → local server (Pythonで作成) に HTTP 通信で再生位置を送信
 - Python (pyglet) で再生位置を更新