第4章　実験

本書では本研究で行った実験について，詳細を述べる．

・仮説

本研究において，アノテーションの数は以下の３つの意味を持つ．(1)1ページあたりのアノテーションの数，(2)アノテーションが必要なページ数，(3)1ページあたりのアノテーションの数とアノテーションの必要なページ数の積

これを踏まえて，本研究では以下の仮説を設定した．

H1 冊子体ボリュームデータからのページ抽出課題において，ページ番号を推定する関数がラプラス方程式を満たすと仮定し，関数のパラメータを最適化すれば，ページ抽出の精度を保ちながら，必要なアノテーション数を減らすことを可能にする．

H1.1 冊子体ボリュームデータからのページ抽出課題において，ページ番号を推定する関数がラプラス方程式を満たすと仮定し，関数のパラメータを最適化すれば，ページ抽出の精度を保ちながら，1ページあたりに必要なアノテーション数を減らすことを可能にする．

H1.2 冊子体ボリュームデータからのページ抽出課題において，ページ番号を推定する関数がラプラス方程式を満たすと仮定し，関数のパラメータを最適化すれば，ページ抽出の精度を保ちながら，アノテーションが必要なページ数を減らすことを可能にする．

H1.3 冊子体ボリュームデータからのページ抽出課題において，ページ番号を推定する関数がラプラス方程式を満たすと仮定し，関数のパラメータを最適化すれば，ページ抽出の精度を保ちながら，1ページあたりに必要なアノテーション数とアノテーションが必要なページ数を同時に減らすことを可能にする．

H1は冊子体ボリュームデータとコンデンサの形状と性質の類似性からコンデンサにおいて電位の初期条件から全体の電位分布を求めるのと同じ手法を用いて，部分的なアノテーションデータという形で与えられるページ番号の初期条件を元に，全体のページ分布を求めることが可能であるという仮説である．

H1.1はラプラス方程式を満たす関数が紙の3次元的な歪みを部分的なアノテーションデータから復元し，ページ全体の形状を推定できるという仮説である．

H1.2はラプラス方程式を満たす関数が，3次元的に歪んだ紙が重なった際に生まれる紙と紙の間隙を紙の形状が与えられた際に推定できるという仮説である．

H1.3はH1.1とH1.2が同時に成立し得るという仮説である．

・実験デザイン

本実験では，第3章において述べた手法によって作成したアノテーションデータを用いる．ただし1ページあたりに得られる最大アノテーション数が図〜に示すようにページによって異なるため，1ページあたりのアノテーション数は15000を上限とした．また実験ごとにランダムに，定められた数のアノテーションデータを取得するようにした．

実験は

1. 1ページあたりのアノテーション数，アノテーションを行うページ番号を設定
2. 設定の条件を満たすアノテーションデータを特定の領域のアノテーションデータに偏らないようにランダムに取得
3. 設定したアノテーションデータを初期条件として，ラプラス方程式を満たす関数を推定
4. 得られた関数から今回対象としている冊子体ボリュームデータのページ分布を算出
5. 算出されたページ分布から初期条件としたアノテーションデータを含む全てのアノテーションデータを算出
6. 人によって作成されたアノテーションデータと，ラプラス方程式によって算出されたアノテーションデータを比較し，本手法を評価

0

20247

20247

1

20093

20093

2

20017

20017

3

19890

19890

4

19788

19788

5

19602

19602

6

19483

19483

7

19508

19508

8

19257

19257

9

19045

19045

[Finished in 73.7s]

working on page0

41613

working on page1

41616

working on page2

41430

working on page3

41430

working on page4

41261

working on page5

41136

working on page6

41261

working on page7

41383

working on page8

41091

working on page9

41694

[Finished in 49.0s]

・検証方法

本研究で設定した仮説について，どのように明らかにするのかを述べる．

H1.1~3を順に検証することで，最終的にH1の妥当性について述べるものとする．

H1.1は，