

B210

文学作品の正規直交化解析に関する一考察

井波 真弓[○](白百合女子大学), 齋藤 兆古(法政大学), 堀井 清之(白百合女子大学)

A Study of Orthonormal Analysis to the Literary Works

Mayumi INAMI, Yoshifuru SAITO and Kiyoshi HORII

ABSTRACT

After applying discrete wavelets multi-resolution analysis to literary works, it is verified that this method considerably removes the analyzed results from the dependency of analyst. Nevertheless, it is difficult to remove the choice and collection of analysis data depending on personal experience exactly. By introducing a general idea of Orthonormal Analysis, duplicated information in analysis data gets objectivity and each data gets same weight to the work. This paper applies Orthonormal Linear Space System Analysis to the selected data comprising the vectors of "The Tale of Genji", after that multi-resolution analysis of discrete wavelets is applied to these vectors. The term vector is built by three keywords of feeling "Inclination", "Disinclination" and "Adoration". As a result, it is revealed that the story is constituted by "Inclination" and "Adoration", and the visualized curve of "Adoration" has common part to "Inclination" in every level in one episode in "The tale of Genji": "Utsusemi story". Further, it is visualized that the same tendency of conventional opinion that the hero's emotion fluctuates due to triflingness of the heartless attitude of the heroin and loss feeling of youth of hero. On the other hand, it is extracted the characteristic more in detail.

Keywords: Orthonormal analysis, Discrete wavelets multi-resolution analysis, Literary works, Independence of an analyst, Detailed analysis,

1. 緒 論

本稿では文学作品に正規直交系線形空間論^{1)・2)}を導入することで、文学作品分析結果に客観的な評価を与え、かつより精緻な分析を行うことを目的とする。文学作品を構成する語彙は個々の語彙が単一の意味を持つとは限らない。このため、文学作品解析において解析対象データの選択と採取はある程度個人的経験に依存せざるを得ない。各章、段落等毎に採取した語彙が構成する物語全体を通じた語彙ベクトルは個人的経験に基づくため必ずしも線形独立とは限らない。このため、語彙ベクトルに対してグラム・シュミットの方法を適用し、語彙ベクトル間の直交性、すなわち、語彙ベクトルの一意性を確保した。さらに、各語彙ベクトル大きさを単位ノルムへ正規化し、物語全体を通して、語彙の選択が与える解析結果への影響を平等化した。このようにして得られた正規直交系ベクトルへウェーブレット多重解像度解析を適用する方法を提案する。

『源氏物語』は十世紀に紫式部によって書かれた世界最古の長篇小説である。「帯木」・「空蟬」・「夕顔」は帯木三帖と言われているが、特に「帯木」の終わりから「空蟬」にかけて内容が連続しており、この二帖が「空蟬物

語」と呼ばれている。「空蟬物語」は青春の「心の惑ひ」と喪失感を源氏に痛感させる物語^{3)・4)}ととらえられている。

2. 解析方法

2.1 解析対象

主人公源氏は桐壺帝の第二皇子で、母、桐壺の更衣と死別後、臣籍に降下し、源氏と名乗った。12歳で元服し、16歳の葵の上と結婚するが、夫婦仲はあまりよくない。後に義理の母である藤壺と密通し、以後女性遍歴をする。理想の女性は桐壺の更衣であって、その影を常に追い求めている。空蟬との出会いは17歳の時である。

空蟬は源氏の恋愛遍歴における最初の女性である。早くに父母を失い、老受領の後妻となり、その後源氏と契ることになる。源氏に撞れつつも拒絶し続け、最後は、小桂を残して去る。

五月雨の降続くある夜、物忌みで宮中の宿直所にこもっている時に貴族たちから未知の中流階級の女性に個性的な魅力のある女性が多いことを聞き、興味を持った。翌日方違えにかこつけて中川にある紀伊の守の屋敷へ行き、空蟬の寝所に忍び込み、強引に契りを結んだ。その後源氏は空蟬に恋文を届けさせるが、空蟬はわが身の立

場をわきまえて返事を出さず、源氏の再訪をも拒否した。源氏は紀伊の守が任国に下った留守の夜、三度中川の屋敷を訪れた。侍女たちが寝静まるのを待って、源氏は空蟬の寝室に向ったが、空蟬は源氏に気づき、小桂を残して抜け出す。源氏は空蟬に逃げられたことに気づいたが、どうすることもできず、残された継娘と契る。二条院に戻った源氏は、小桂をいつも肌身離さず、その人の香を懐かしく思うのであった。

Table 1 Selected Element

要素	事例
第1要素「会う方向」	相手に会いたい と思う気持ち
第2要素「会わない方向」	相手に会うまい と思う気持ち
第3要素「思慕」	ひたすら相手を 思う気持ち

Table 2 Examples of element in Genji

要素	事例
「会う方向」 (Inclination)	・さる心して人とかく静めて御消息あれど (源氏の君は、そのつもりでお供の人々を早くから寝静まらせて)
「会わない方向」 (Disinclination)	かの人々の心を爪はじきをしつつ恨みたまふ (あの女の人の心根をつまはじきにしてはお恨みになる)
「思慕」 (Adoration)	・空蟬の身をかへてける木のものもとなほ人がらのなつかしきかな (せみが抜け殻を残して姿を変え、去ってしまった後の木の下で、も抜けの殻の衣を残していったあの人の気配をやあり懐かしく思っている)

2.2 正規直交解析

線形とはデータの重ねが成り立つということであり、重複する要素を含むデータは線形独立なベクトルを構成しない。

文学作品にベクトルの概念を用いる。本稿では、作品の各段落ごとの特定語彙、ここでは感情を表す語彙数を要素とするベクトルを考える。作品の構成を継時的に考察するために源氏について語られている部分から「会う方向」「会わない方向」「思慕」を示すことばを要素に選び、段落ごとの使用頻度を調べる。Table 1, 2は要素を示す。ただし、和歌の場合は1首であっても1段落とした。これらの各語彙が構成するベクトルをいわゆる線形空間のベクトルと対応させる。

次は、内積の概念である。この概念はデータの積和、すなわち線形空間における内積と等価である。それぞれの要素の内積は

$$\text{Inclination}^T \cdot \text{Disinclination} \neq 0$$

$$\text{Inclination}^T \cdot \text{Adoration} \neq 0$$

$$\text{Disinclination}^T \cdot \text{Adoration} \neq 0 \quad (1)$$

となるため、重複する要素を含んでおり、線形独立ではない。

各語彙が構成するベクトル間で角度の概念が成り立つ。ベクトル間の角度の概念は、同一要素を含む割合を角度で表現する方法で、文学作品において語彙ベクトル間の重複性に対応する。0度は全く同じ方向、性質のベクトル、90度は線形独立なベクトルで、重複要素がない。180度は全く相反する方向、性質のベクトル語彙で、ベクトル間の角度が90度に近い場合は両ベクトルが独立して作品を構成していることを意味する。

本論文においてはベクトルをグラム・シュミットの方法によってすべて直交化する。直交化によって語彙ベクトル間の重複要素が除かれる。さらに直交化されたベクトルを単位ノルムに正規化することで作品を構成するベクトルの重みを均一化する。ベクトル間を直交化し、さらにそれぞれのベクトルのノルムを1へ正規化した系を正規直交系という。

作品を構成する要素は「会う方向」で表現されたものであっても、ある場合には「会わない方向や「思慕」の気持ちを含むこともある。この意味で、「会う方向」、「会わない方向」、「思慕」ベクトルで、ベクトル間に重複する要素が存在するのは当然である。この重複度を表す指標としてベクトル間の角度やパターン一致度を調べた。その結果、「会う方向」ベクトルと「会わない方向ベクトル間は 40.332 [Deg.]

$$\cos^{-1} \left[\frac{\text{Inclination}^T \text{Disinclination}}{\|\text{Inclination}\| \|\text{Disinclination}\|} \right] \times \frac{180^\circ}{\pi} = 40.332 [\text{Deg.}] \quad (2)$$

で要素の重複度が比較的高い。「会う方向」ベクトルと「思慕」ベクトル間は 55.2613 [Deg.]、「会わない方向ベクトルと「思慕」ベクトル間は 57.6793 [Deg.]で、要素の重複度は比較的低いことがわかった。ここでは、各感情ベクトルで重複する要素を取り除き、それぞれのベクトルが重複する要素を含まない互いに独立な直交したベクトル成分だけを抽出する。

「会う方向」ベクトルを基準ベクトルとして、「空蟬物語」に登場する感情ベクトルをグラム・シュミットの方法で互いに直交化する。最初は、「会わない方向ベクトルから「会う方向」ベクトルに平行する成分を抽出する。

$$\text{DisinclinationComp}$$

$$= \text{Disinclination} - \text{Inclination} \times \frac{\text{Inclination}^T \text{Disinclination}}{\|\text{Inclination}\| \|\text{Disinclination}\|} \quad (3)$$

ベクトル Disinclination Comp と Inclination 間の内積は

$$\text{Disinclination}^T \cdot \text{Inclination} = 0 \quad (4)$$

となり、互いに直交し重複する要素を含まない。

「会う方向」ベクトルを基準ベクトルとした正規直交系の章毎のベクトルは以下の様になる。

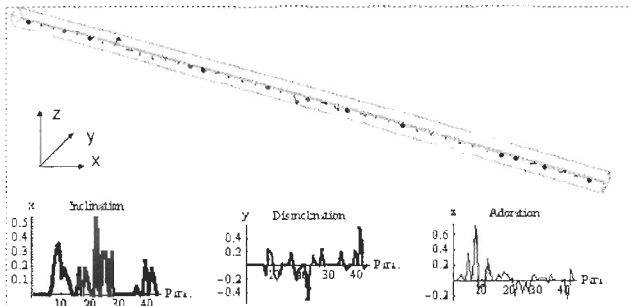


Fig. 1 Vectored view of Orthnormal Analysis : sentimental pattern of Genji

x:Inclination, y:Disinclination, z:Adoration

正規直交化されたデータにウェーブレット多重解像度解析を適用する。

多くの離散値系ウェーブレット変換基底が提案されているが、最も簡明な Haar (ハール) 基底は解析結果の解釈が明解であるため、これを採用する。Haar 基底は前半、後半という考え方でソーティングできるため、各種の解析の意味が把握しやすい側面がある。ウェーブレット多重解像度解析で、レベル 1 は全データの平均値に相当し、レベル 2 は全データを半分に分けたときの変化率、レベル 3 は全データを 4 分割したときの変化率、レベル 4 は全データを 8 分割したときの変化率、レベル 5 は全データを 16 分割したときの変化率、最も高次のレベルは隣接する要素間の変化率をそれぞれ与える。

離散値系ウェーブレット多重解像度解析結果を Fig. 3 と Fig. 4 に示す。実際のデータ数としての段落は 43 章までであるが、解析には 2 のべき乗のデータが必要であるため最後の段落に 44 から 64 段落をゼロデータとした。また、Fig. 3 と Fig. 4 の結果はゼロを追加した段落を削除してある。従来の解析結果と比較することで、文学作品に対するウェーブレット多重解像度解析は、比較的高次の多重解像度解析結果が作品を書く筆者の気持ちや感情的ゆらぎなどを表す考察対象として意義があることを示す。

データベクトルを Y 、ウェーブレット変換行列を W とすればウェーブレットスペクトラム S は次式で与えられる。

$$S = WY \quad (5)$$

ウェーブレット多重解像度解析で、レベル 1 はスペクトラム行列 S の第 1 要素のみを残し他の要素をゼロとしてウェーブレット逆変換式 (6) で得られる。

$$S' = \begin{bmatrix} S_0 \\ \cdot \\ \cdot \\ S_n \end{bmatrix}, \quad D_0 = W^T \cdot S' \quad (6)$$

3. 結果と考察

「空蟬物語」における源氏と空蟬に関する Fig. 2 から Fig. 7 の解析結果から、一度の逢瀬の後、相手に惹かれながらも会うことがなかったが、源氏は思慕の情を募らせていったことが示された。

3.1 解析結果

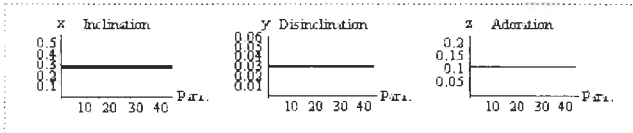


Fig. 2 Level 1 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

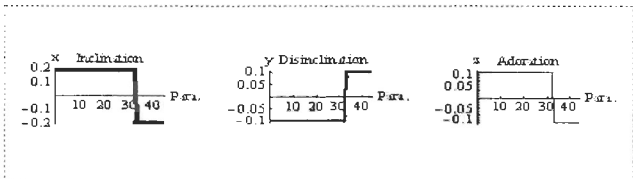


Fig. 3 Level 2 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

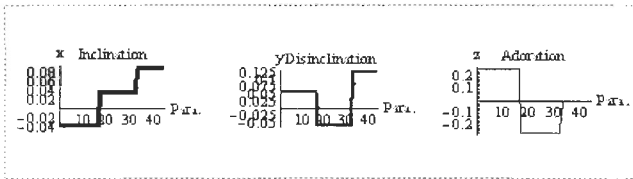


Fig. 4 Level 3 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

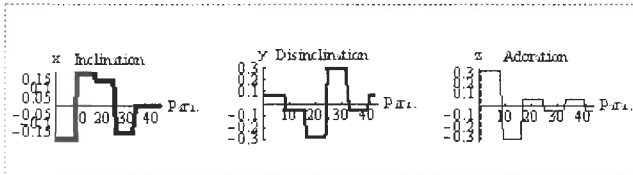


Fig. 5 Level 4 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

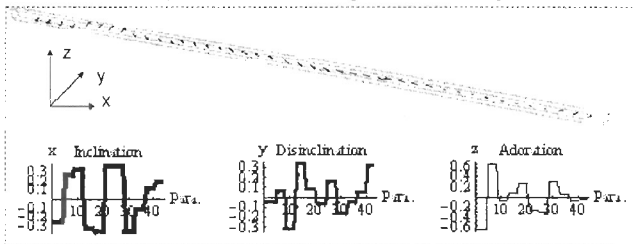


Fig. 6 Level 5 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

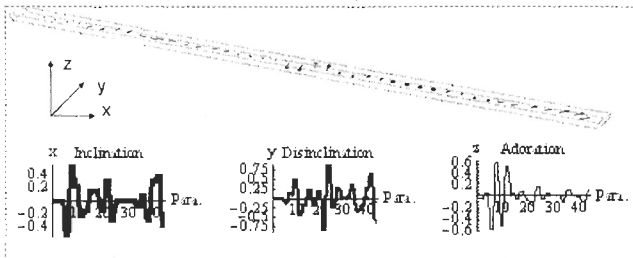


Fig. 7 Level 6 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

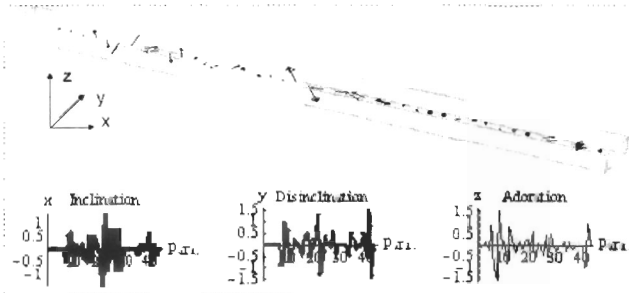


Fig. 8 Level 7 of the discrete wavelets multi-resolution analysis : sentimental pattern of Genji

Fig. 2 のレベル 1 は作品全体の平均的な感情用語の頻出度合いを表す。「空蟬物語」では、作品全体として「会う方向」と思慕の大きさが構成されていることがわかる。

Fig. 3 のレベル 2 は「源氏物語」を 2 分割した場合の感情語彙の頻出度合いを表す。ほとんどの段落は「会う方向」が支配的であるが、「会わない方向」と思慕が互いに反転した構成になっている。また、「会う方向」と思慕が同一傾向を持つ事が興味深い。

Fig. 4 のレベル 3 は「源氏物語」を 4 分割した場合、すなわち 16 章ごとに平均化された情報の感情語彙の頻出度合いを表す。ベクトルの成分として思慕の頻出度合いが支配的であり、最初の第 16 章までは頻出度合いが高いが、第 32 章までは減少し、最終章にかけて平均的な頻出度合いである。「会う方向」と「会わない方向」の頻出度合いはほぼ単調に増加し、思慕は初期は大きい後半で減少する。

Fig. 5 のレベル 4 は「空蟬物語」を 8 分割した 8 章ごとに平均化された情報である。ベクトルの成分として思慕の頻出度合いが支配的であり、最初の第 1-8 章と第 9-16 章では大きく増減するが、その後は平均的で増減が低下する。「会わない方向」は第 17-24 章と第 25-32 章で頻出度合いが大きく増減している。「会う方向」は「会わない方向」が減少している時には対照的に頻出度合いが増加している。

Fig. 6 のレベル 5 は「空蟬物語」を 16 分割した場合の 4 章ごとに平均化された情報である。ベクトルの成分として思慕の頻出度合いが支配的であり、最初の第 1-4 章と第 5-8 章では大きく増減するがその後小さな増減を繰り返しながら平均的となる。「会う方向」と「会わない方向」は頻出度合いが対照的に推移するが、後半は両者とも増加している。

Fig. 7 のレベル 6 は「空蟬物語」を 32 分割した場合の 2 章ごとに平均化された情報である。ベクトルの成分として「会わない方向」の頻出度合いが支配的である。思慕は最初では頻出度合いが大きく増減していたが、その後平均的な頻出度合いとなっている。

Fig. 8 の最も高次のレベル 7 は各章ごとの感情語彙の頻出度合いであり、「会わない方向」思慕の度合いが支配的であるがレベル 7 はデータの平均化がなされていない

ため、隣接するデータ間のバラツキを強調した結果を含んでいる。

Figs. 6, 7, 8 のベクトル図はそれぞれ構成がよく把握されるレベル 4, 5, 6 のみを提示した。特にレベル 6 は明らかに空蟬物語の山場が中心部の章に在ることを意味している。

4. 結論

- 1) 作品全体が「会う方向」と「思慕」の大きさが構成されている。
- 2) レベル全体を通じた「会う方向」「会わない方向」の変化に比べて「思慕」は変化が少なく、ドミナントである。
- 3) レベルの違いがあっても、「思慕」の描く可視化曲線には「会う方向」ベクトルと共通部分があり、かつ「空蟬物語」における暗黙知が抽出された。
- 4) 「空蟬物語」は源氏に青春の「心の惑ひ」と喪失感を痛感させる物語であるという従来の説と同じ傾向が可視化され、かつ、特色がより詳細に検証された。
- 5) 離散値系ウェーブレット変換の多重解像度解析は各レベルの解析結果に着目することでより精緻な解析と解釈を可能にした。
- 6) 理工学は平均値が重要な世界であるが、文学においては、言語によって人間の心のゆらぎが表現される。したがって、低次だけでなく高周波の部分が無視できないことが考察された。

参考文献

- 1) 齋藤兆古：ウェーブレット変換の基礎と応用—Mathematica で学ぶ、朝倉書店（1998）p. 39, pp. 93-95.
- 2) 堀井清之、齋藤兆古：特許「文学作品解析方法および解析装置」、特願 JP10-102673A.
- 3) 藤田加代：空蟬——人物造型の特異性——、源氏物語講座第二巻 物語を織りなす人々、(株) 勉誠社（1991）p. 80.
- 4) 井波 真弓、岩崎 晴美、宮沢 賢治、土屋 宏之、齋藤兆古、堀井 清之”『源氏物語』における源氏と空蟬の恋—ウェーブレット多重解像度解析—、可視化情報学会論文集、Vol. 25, No. 5（2005）pp. 8-12.