

博士論文

機械学習による
紙特性管理手法の高度化に関する研究
**A study on Improvement of Paper Property Management Method
Using Machine Learning**

国立大学法人 横浜国立大学

大学院環境情報学府

石川 卓

Taku Ishikawa

2015年3月

あらまし

経済活動において銀行券，諸証券，旅券および ID 文書などの物理媒体（以下，貴重印刷物という。）は重要な役割を果たしている。貴重印刷物は通常の印刷物とは異なり，それぞれ特殊な価値や役割を持つ。このため，貴重印刷物の偽造およびその使用は，法律（通貨偽造・変造罪，偽造通貨の行使罪）によって固く禁じられている。しかし偽造は後を絶たず，古くから現代に至るまで常に社会的な問題となっている。偽造の脅威に対抗すべく，貴重印刷物の偽造を防止するために，製造者は様々な偽造防止対策を施しているが，世の中の技術の進歩に連動して，偽造者が得られる技術は向上する。このため，貴重印刷物の偽造防止効果の向上は常に求められている。貴重印刷物の偽造防止効果の向上においては，特殊な印刷模様，磁気特性，特殊な材料の付与など様々な技術的な要素が検討されている。また一方で，標準や評価基準を導入し，製品の品質を高い水準で維持する取り組みも重要である。この取り組みはシンプルだが真正品が持つ特性のバラツキを小さくすることに繋がり，偽造品と真正品の区別を容易にすることから高い効果が期待できる。近年では偽造防止技術を対象とする ISO 技術委員会が発足され，欧州標準化委員会（The European Committee for Standardization; CEN）による貴重印刷物の製造管理体制における安全標準の開発も行われている。このような背景から，貴重印刷物の特性の管理は重要とされている。そして日本国内においては，代表的な貴重印刷物である銀行券や旅券の基材に紙媒体が用いられている。貴重印刷物に用いられている紙には通常の紙とは異なる特性が求められる。例えば，市中における流通に耐えうる耐久性，特殊な偽造防止技術の効果を安定して発現させる特性，そして機械読取に対する安定性などが挙げられる。これらの要求を満たすために，貴重印刷物の基材として用いられる紙には，通常の紙と比較して，高レベルな品質の評価および管理が必要とされる。このような背景の中，紙の製造工程において，紙や紙料の品質や特性を管理するため，それら进行评估する試験方法が規格化されている。さらに，より正確な評価を行うための研究が行われ，高精度な測定装置やその測定データから多様な特徴量を算出する方法が開発された。その結果，従来と比較して製紙工程において多様な情報が取得可能となった。しかし，紙の特性においては繊維の特性，薬品の特性，製造条件の変動など，その特性に影響する要因は多岐に渡る。また，それぞれが相互に影響し合うため，その物理現象は非常に複雑であり，現在においても紙の特性の詳細な評価，そして紙が構築されるメカニズムを把握することは難しい課題であり，より正確に紙の特性を管理することが可能な手段が求められている。従来，このような課題においてはデータ間の相関係数の参照や主成分分析（Principal Component Analysis; PCA），重回帰分析（Multiple Liner Regression; MLR）などの統計手法が多く用いられている。統計的なアプローチは手法としてよく確立され，実用的な問題への適用が容易である。また，その解析結果もシンプ

ルで、結果と要因の関係の説明が容易である。上記のような特徴から、統計的手法は広い分野において使用されている。しかしながら、このような統計手法は複雑な問題への適用において、精度を向上させることが容易ではないという問題がある。一方、近年ではパーソナルコンピュータの進歩に伴い、進化計算や神経回路網 (Neural Network; NN), サポートベクターマシン (Support Vector Machine; SVM) などを代表とした機械学習によって大規模データを解析する手法が活発に研究されている。機械学習はコンピュータの計算処理によって自動的に情報の探索 (学習) を行い、目的の処理を行う情報処理 (モデル) を構築する。学習を通じて自動的に大規模データの中から、人が気づくことが難しい基準やルールを発見し、複雑なモデルを柔軟に構築可能なため、様々な問題に適用可能である。紙の製造に関する従来研究として、紙の特性の解析における実問題へ機械学習を適用した例もある。しかし、一般に機械学習は自由度の高いモデルを構築可能である反面、一方でその構造が複雑になり、結果と要因の関係の説明が困難となる。このため、実問題への適用において、機械学習によって構築されたモデルが示す結果の信頼性を保障および説明することが難しいという問題がある。具体的には製品ラインにおける検査や貴重印刷物の真偽判別において機械学習によるモデルを使用すると、機械学習のモデルを使って得られた結果について、その結果に至った根拠や理由を求められる。特に、構築されたモデルがイレギュラーな結果を示した場合には、その原因の特定は強く求められる。このため、モデルの持つ結果の正確さとともにその結果に至る過程の説明も重要であり、製造工程で実際に運用する際には必要不可欠である。

そこで本論文では、紙の製造において取得可能となった多様かつ大量の情報を活用する方法として機械学習を用い、紙の特性をより高精度に分類、予測することが可能でかつ、人に理解し易い紙の分類および特性予測モデルを構築することを目指した。

まず、製造された紙の分類器を特徴量選択型 SVM によって構築する方法を提案する。実験では市販の印刷用紙および叩解回数、薬品、抄紙ワイヤーの形状など製造条件を変更した手すきシートをサンプルとし、光学スキャナによって撮像した各サンプルの透過画像における複数の画像特徴量から各サンプルを分類する分類器を構築し、その分類精度および特徴量の取捨選択において、提案方式の有効性を示す。

次に、進化型条件判断ネットワーク (Evolutionary Decision Network; EDEN) を用いて紙の分類器を構築する方法を提案する。EDEN の構造においてはノードの遺伝子型を拡張し、単一の特徴量間における特徴量間の相互作用の表現を図る。ノードを拡張した EDEN が多様な条件下において製造された紙の分類において、高い分類性能と、高い可読性を持った分類器を構築可能であることを示す。

そして、Cartesian Genetic Programming を用いることによって紙料中の繊維の特徴量から紙の光学特性を予測するモデルを構築する方法を提案する。CGP においては実数部をノードの遺伝子型へ追加し、その自由度を向上させることによって予測モデルのコンパクト化を図り、予測精度と構造のシンプルさの両立において提案方式が有効であることを示す。

Abstract

Security printings such as banknotes, bonds, passport have the importance to our economy and the certification of personal rights. There are many kinds of security printings, and each of them have the special role. Therefore, counterfeiting of security printings is strictly prohibited by law. However, the crimes of counterfeiting are never ceasing and have become a social problem since a long time ago. Security printing producers have imparted various anti-counterfeiting technologies to prevent counterfeiting. However, the technologies which give to counterfeiter are increasing in conjunction to technical advance. Therefore, the improvement of anti-counterfeiting effect of security printings is always required. Several methods such as a unique printing pattern, magnetic feature and special materials are examined for the improving of anti-counterfeiting effect in security printings. In this problem, the introduction of evaluation criteria and standard and the maintaining of high levels of product quality can be mentioned as one of the means. This approach is simple; however, it can be expected to reduce the dispersion of property of genuine products, therefore its effectiveness is high. In recent years, ISO technical committee for anti-counterfeiting technology has been established and CEN (the European Committee for Standardization) have developed the safety standard in production management of security printings. In this background, the management of property of security printings is important. Currently, Paper is often employed as substrate of typical security printings such banknotes and passport. The paper employed for security printings is required the special characteristics that are different from ordinary papers. For example, the durability withstands the distribution in the society, the characteristic for expressing the effect of anti-counterfeiting technologies stably and the stability for machine reading can be mentioned. Therefore, high levels of production quality and management is required to the paper employed for security printings compared to ordinary papers. In this background, the measurement method which evaluating the paper properties has been standardized for managing them in the paper production line. Moreover, high accuracy measuring device and the method which can calculate several features from measurement data have been developed for accurate evaluation. As a result, various kinds of information have been available in the paper production compared to the past. However, the effect factors which affect the paper properties are diverse such as fiber feature, chemicals, production condition. Moreover, each factor has the mutual effect to each other; therefore the accurate evaluation of paper and the grasping of paper formation mechanism are still difficult theme. Therefore, the method which can accurately manage the paper properties is required. Generally, statistical methods such as a correlation coefficient, Principal Component Analysis (PCA) and Multiple Liner Regression (MLR) are employed for these problems. Statistical approach is well

established, and easy to apply for practical problems. Moreover, its analysis result is simple, therefore explain of relationship between result and factors is easy. From the features described above, statistical methods are employed in various fields. However, statistical methods are difficult to apply to complicated problems. On the other hand, Machine Learning methods such as Evolutionary Computation, Neural Network, and Support Vector Machine have been actively studied to analyze large scale data in conjunction to improvement of personal computer. Machine learning method automatically finds the information with computation and creates the model which operates the processing of various purposes. Machine learning can find the standards and rules in large scale data which are difficult to find by human, and create complicated models automatically and flexibly. Therefore, it can be applied various problems. There are some previous studies of applying machine learning to practical problems of analyzing paper properties. However, the structures of models created with machine learning are often too complicated in explain the relationship between result and factors. Because of this, there is a problem to keep the reliability for the results which are shown by models created with machine learning. Specifically, the basis and reason of deriving result are often required in the inspection process and authenticity discriminating. Especially, it is strictly required if the model showed the irregular result. Therefore, explanations of process of deriving the result are important equally to accuracy of result when employing the method in the practical production line. In this thesis, machine learning methods are employed to create the classifier and prediction model which has high accuracy and readability using large scale data in paper production and paper properties. First, I propose a method for creating the paper classifier with feature selection type SVM (FS-SVM) and show the effectiveness of method. In the experiment, classifier models are created for various papers with image features calculated from transparent image of paper and evaluate the classification accuracy and selection of features. Then, I proposed a method for creating the paper classifier with Evolutionary Decision Network (EDEN) and show the effectiveness of method. About the structure of EDEN, I expanded the node for expressing the mutual effect between features in simplex feature group. Finally, I propose a method for creating the paper property prediction model with Cartesian Genetic Programming (CGP) and show the effectiveness of method. About the structure of CGP, I expanded the node for improving the flexibility for improving the simplicity of the model.

目次

第1章	序論	1
1.1	背景	1
1.2	本論文の構成	2
1.3	貴重印刷物とその偽造	4
1.4	紙と貴重印刷物	5
1.5	貴重印刷物用紙の製造と品質管理の課題	9
1.6	研究の目的	12
第2章	紙と紙料の評価方法	13
2.1	製紙に関わる一般論	13
2.1.1	紙の定義	13
2.1.2	紙の分類	13
2.1.3	紙の材料	14
2.1.4	紙の基本的性質	15
2.1.5	紙の特性を表す指標	15
2.1.6	製紙用パルプの特性を表す指標	17
2.2	紙の評価技術およびその研究	19
2.2.1	紙の評価方法	19
2.2.2	紙の評価に関する従来研究	20
2.3	製紙用パルプの評価技術およびその研究	22
2.3.1	繊維特性の測定方法	22
2.3.2	紙料の評価技術に関する従来研究	23
2.4	評価用手すきシートの作製方法	26
第3章	本研究に関する機械学習手法	28
3.1	機械学習 (Machine Learning)	28
3.2	進化計算 (Evolutionary Computation; EC)	28
3.2.1	進化戦略 (Evolutionary Strategy; ES)	29
3.2.2	遺伝的アルゴリズム (Genetic Algorithm; GA)	29
3.2.3	遺伝的プログラミング (Genetic Programming; GP)	31
3.3	ニューラルネットワーク (Neural Network; NN)	33
3.4	サポートベクターマシン (Support Vector Machine; SVM)	34
3.5	機械学習を用いた紙の製造に関する従来研究	35
第4章	特徴量選択型 SVM による紙の分類器の自動構築	37

4.1	はじめに	37
4.2	特徴量選択型 SVM (Feature selection type SVM; FS-SVM)	38
4.3	実験	40
4.3.1	テストシート.....	40
4.3.2	テストシートの撮像方法.....	41
4.3.4	画像特徴量の算出.....	43
4.3.5	特徴量選択型 SVM に関するパラメータ	44
4.3.6	分類対象の設定.....	45
4.3.7	比較手法	48
4.4	結果	49
4.4.1	各分類器の分類正答率.....	49
4.4.2	選択された特徴量.....	49
4.5	考察	52
4.5.1	従来手法との比較.....	52
4.5.2	てん料の添加率の異なる紙の分類.....	52
4.5.3	分類に有効な特徴量.....	53
4.5.4	特徴量の追加.....	53
4.6	まとめ	54
第5章	Evolutionary Decision Network を用いた紙の分類器の自動構築	55
5.1	はじめに	55
5.2	Evolutionary Decision Network (EDEN)	56
5.2.1	EDEN の概要.....	56
5.2.2	EDEN の基本構造.....	56
5.2.3	EDEN における相互作用の表現.....	57
5.2.4	EDEN における遺伝子操作.....	58
5.2.5	EDEN に使用するノードの拡張.....	59
5.2.6	ノードを拡張した EDEN の検証.....	61
5.3	実験	64
5.3.1	テストシート.....	64
5.3.2	画像特徴量の算出.....	64
5.3.3	分類対象	64
5.3.4	EDEN の学習パラメータ	65
5.4	結果	66
5.4.1	分類結果	66
5.4.2	構成された分類器.....	67
5.5	考察	68

5.5.1	他の手法との分類正答率の比較.....	68
5.5.2	分類器の構造について.....	68
5.6	まとめ.....	70
第6章	Cartesian Genetic Programming を用いた紙の特性予測器の自動構築	71
6.1	はじめに.....	71
6.2	Cartesian Genetic Programming (CGP)	72
6.2.1	CGP の概要.....	72
6.2.2	CGP の構造の検討および拡張.....	74
6.2.3	拡張 CGP の検証.....	76
6.3	実験.....	77
6.3.1	テストサンプル.....	77
6.3.2	繊維特徴量.....	78
6.3.3	紙の光学特性.....	80
6.3.4	比較手法	80
6.3.5	機械学習に関するパラメータおよび条件.....	81
6.3.6	予測モデルの検証方法.....	83
6.4	結果.....	84
6.4.1	CGP の学習における適応度の推移.....	84
6.4.2	予測精度の比較.....	85
6.4.3	次元数の削減の比較.....	86
6.4.4	CGP が構築した予測モデルの一例.....	87
6.5	考察.....	88
6.5.1	各モデルの予測性能.....	88
6.5.2	紙の特性へ影響を与えるその他の要因.....	88
6.5.3	CGP によって選択された特徴量.....	88
6.5.4	CGP によって構築された予測モデルについて	89
6.6	まとめ.....	90
第7章	総括	91
7.1	本論文で得られた成果.....	91
7.2	今後の課題	92
謝辞	93	
参考文献	94	
研究業績	100	

図目次

図 1.1	本論文の構成.....	3
図 1.2	偽造日本銀行券の発見枚数の推移.....	4
図 1.3	第1 認証技術の例.....	6
図 1.4	第2 認証技術の例.....	6
図 1.5	日本銀行券の白黒すき入れ.....	8
図 1.6	紙の地合の違い.....	8
図 1.7	国内における自販機台数の推移.....	10
図 2.1	酸性紙の劣化.....	14
図 2.2	日本銀行券の原料.....	14
図 2.3	紙における繊維の状態.....	15
図 3.1	進化的計算法のグループ.....	28
図 3.2	基本的な GA の処理の流れ.....	31
図 3.3	GP の表現型の例.....	32
図 3.4	NN の構造.....	33
図 3.5	SVM におけるマージン.....	34
図 4.1	特徴量選択型 SVM における最適化手順.....	39
図 4.2	特徴量選択型 SVM における遺伝子型.....	39
図 4.3	赤外透過画像撮像装置.....	42
図 4.4	各サンプルの赤外透過画像撮像.....	42
図 4.5	ワイヤーの特徴が現れている紙の PS 画像の一例.....	48
図 4.6	FS-SVM の学習における最良個体の推移.....	50
図 5.1	EDEN の構造の例.....	56
図 5.2	EDEN における異なるデータ群間の相互作用の例.....	57
図 5.3	EDEN の遺伝子型の例.....	59
図 5.4	EDEN において拡張したノードの例.....	61
図 5.5	EDEN において拡張したノードの遺伝子型の例.....	61
図 5.6	EDEN によって構築された分類器モデルの一例.....	67
図 5.7	各手法の分類正答率の比較.....	70
図 6.1	基本的な CGP の表現型の例.....	73
図 6.2	本研究で拡張した CGP の遺伝子型と表現型.....	76
図 6.3	CGP と拡張 CGP の予測性能の比較.....	76
図 6.4	作製した手すきシートの一例.....	77

図 6.5	繊維特徴量測定機.....	78
図 6.6	繊維特徴量測定器によって撮像した繊維画像の一例.....	79
図 6.7	各テストシートの画像の一例.....	80
図 6.8	CGP における適応度の推移.....	84
図 6.9	各手法によって構築されたモデルの予測精度の比較.....	85
図 6.10	予測モデルに使用された特徴量数の比較.....	86
図 6.11	CGP によって構成されたモデルの例.....	87
図 7.1	原料から製造工程までのすべての要素を考慮した予測システム.....	92

表目次

表 1.1	日本銀行券の製紙工程.....	11
表 2.1	実験用手すきシートの作製手順の一例.....	27
表 4.1	手すきシートサンプルの作製条件.....	41
表 4.2	画像特徴量一覧.....	44
表 4.3	特徴量選択型 SVM の学習パラメータ	45
表 4.4	各 Test における分類対象.....	47
表 4.5	各 Test で用いた画像枚数.....	47
表 4.6	各分類器における正答の定義.....	48
表 4.7	FS-SVM による各 Test における分類正答率.....	50
表 4.8	FS-SVM によって各分類器に選択された特徴量	51
表 4.9	フーリエ変換と相互相関法による紙の異同識別法による分類正答率	52
表 5.1	EDEN および拡張 EDEN の学習パラメータの設定.....	62
表 5.2	EDEN と拡張 EDEN によって構築された各モデルの分類性能の比較.....	63
表 5.3	EDEN と拡張 EDEN の適応度の比較	63
表 5.4	EDEN の学習パラメータ設定	65
表 5.5	EDEN が構築した分類器による各 Test における分類正答率.....	66
表 6.1	CGP および拡張 CGP のモデルにおいて使用されたノード数および特徴量数.....	76
表 6.2	繊維特徴量のリスト.....	77
表 6.3	繊維特徴量測定器が出力する繊維特徴量.....	79
表 6.4	各機械学習手法のパラメータ	82
表 6.5	CGP に使用した計算ノード.....	83
表 6.6	各繊維特徴量の選択頻度.....	88

第1章 序論

1.1 背景

経済活動において銀行券、諸証券、旅券およびID文書などの物理媒体（以下、貴重印刷物という。）は重要な役割を果たしている。貴重印刷物は通常の印刷物とは異なり、それぞれ特殊な価値および役割を持つ。このため、貴重印刷物の偽造およびその使用は、法律（通貨偽造・変造罪、偽造通貨の行使罪）によって固く禁じられている。しかし偽造は後を絶たず、古くから現代に至るまで常に社会的な問題となっている [1]。偽造の脅威に対抗すべく、貴重印刷物の偽造を防止するために、製造者は様々な偽造防止対策を施しているが、世の中の技術の進歩に連動して、偽造者が得られる技術は向上する。このため、貴重印刷物の偽造防止効果の向上は常に求められている。貴重印刷物の偽造防止効果の向上においては、特殊な印刷模様、磁気特性、特殊な材料の付与など様々な技術的な要素が検討されている。また一方で、標準や評価基準を導入し、製品の品質を高い水準で維持する取り組みも重要である。このような取り組みはシンプルだが真正品が持つ特性のバラツキを小さくすることに繋がり、偽造品と真正品の区別を容易にすることから高い効果が期待できる。近年では偽造防止技術を対象とするISO技術委員会が発足され、欧州標準化委員会（The European Committee for Standardization; CEN）による貴重印刷物の製造管理体制における安全標準の開発も行われている [2]。このような背景から、貴重印刷物の特性の管理は重要とされている。現在、代表的な貴重印刷物である銀行券や旅券の基材としては紙媒体が多く用いられている。このような貴重印刷物に用いられている紙には通常の紙とは異なる特性が求められる。例えば、市中における流通に耐えうる耐久性、特殊な偽造防止技術の効果を安定して発現させる特性、そして機械読取に対する安定性などが挙げられる。これらの要求を満たすために、貴重印刷物の基材として用いられる紙には、通常の紙と比較して、高レベルな品質の評価および管理が必要とされる。このような背景の中、紙の製造工程において、紙や紙料の品質や特性を管理するため、それら进行评估する試験方法が規格化されている。さらに、より正確な評価を行うための研究が行われ、高精度な測定装置やその測定データから多様な特徴量を算出する方法が開発された。その結果、従来と比較して製紙工程において多様な情報が取得可能となった。しかし、紙の特性においては、繊維の特性、薬品の特性、製造条件の変動など、その特性に影響する要因は多岐に渡る。また、それぞれが相互に影響し合うため、その物理現象は非常に複雑であり、現在においても紙の特性の詳細な評価、紙が構築されるメカニズムを正確に把握することは難しい課題であり、より正確に紙の特性を管理することが可能な手段が求められている。

1.2 本論文の構成

本論文の構成を以下に示す.

第 1 章では, まず貴重印刷物における偽造の実態を示す. 次に貴重印刷物と紙の関係について述べる. そして, 貴重印刷物用紙の製造工程として日本銀行券の製紙工程の概要について述べ, その特性管理における課題について述べ, 最後に本研究の目的を述べる.

第 2 章では, 本研究に関わる製紙の一般論, 紙および紙料の評価技術について述べる. まず製紙の一般論として, 紙の定義や材料, 性質および特性などについて述べ, 次に紙および紙料の試験方法および各技術に関する従来研究について述べる. また, 実験に使用する評価用紙の作製方法についても述べる.

第 3 章では, 本研究に関する機械学習手法の概要および機械学習を用いた従来研究について述べる.

第 4 章および第 5 章においては紙の分類に関する研究, 第 6 章においては紙の特性の予測に関する研究について述べる.

第 4 章では, 特徴量選択型 SVM を用いた紙の分類器の自動構築実験について説明する. 本章では大量の特徴量の中から必要な特徴量のみを自動選択する手法を用いることによって, 人が納得する分類器を構築することに重点を置いた. 具体的には紙の透過画像から算出可能な複数の画像特徴量から特徴量選択型 SVM によって分類器を構築し, その分類性能および分類に有効な画像特徴量について検証を行う.

第 5 章では, 紙の分類問題において分類過程に分かり易さに重点を置き, 人間に分かり易い分類器のモデルを構築するため, Evolutionary Decision Network (EDEN) を用いた紙の分類器の自動構築実験を行う. 紙の透過画像の複数の画像特徴量から分類器を構築し, その分類性能および分類過程について検証を行う.

第 6 章では, 紙の特性を紙料の特性から予測するモデルを Cartesian Genetic Programming (CGP) を用いて自動構築する実験について説明する. 繊維の複数の特徴量から紙の特性を予測する予測器を構築し, その予測性能を従来手法と比較することによって検証する. また, 構築された予測モデルにおける予測の過程について確認する.

最後に第 7 章にて本論文のまとめと今後の課題について述べる.

図 1.1 に本論文の構成と公表論文との対応を示す. なお, 文献番号[Pub.1]~[Pub.4]は p.96 に記載の研究実績による.

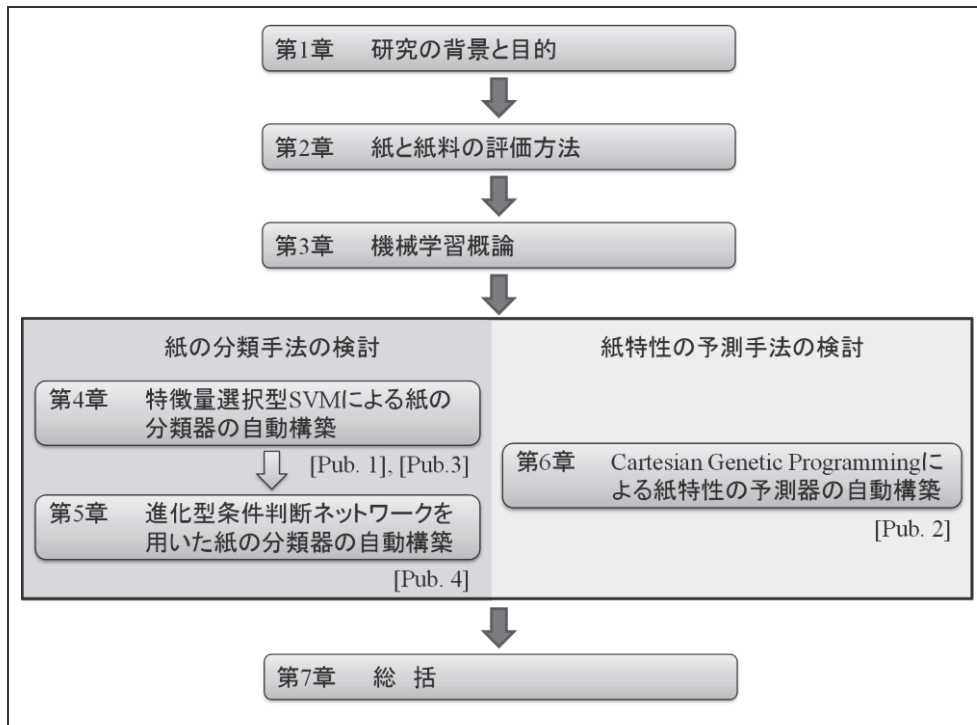


図 1.1 本論文の構成

Fig. 1.1 Structure of this thesis

1.3 貴重印刷物とその偽造

本節では貴重印刷物の代表格である銀行券とその偽造について述べる。日本国では銀行券は経済活動において非常に重要な役割を果たしている。家庭用パーソナルコンピュータの普及に加え、スマートフォン、タブレット端末の普及台数が急激に増加し、インターネットなどの情報通信ネットワークによる収納サービスなどの利用実績は2005年には200万件に満たなかったのが、2012年には1400万件を超えた [3]。しかし、日常的な支払いにおける主な決済手段における現金の使用率は、千円から五千円の間では92%、五千円から一万円の間で84.3%であり、依然として第一の決済手段となっている [4]。以上からIT技術の普及によって決済手段が多様化しているが、手軽さや融通性などの理由から銀行券が主流な取引手段として利用されていることがわかる。

貴重印刷物の偽造の事例として、日本銀行券の偽造の発見枚数 [5] [6]を図1.2に示す。千円券から一万円券の4券種の偽造発見枚数の合計が2000年から平成2005年にかけて急激に増加していることが確認できる。この要因として、パソコン、プリンタ、スキャナなどDTP (Desk Top Publishing)と呼ばれる画像処理機器および印刷装置に関する技術が発展することによって、高精度な印刷環境が一般家庭でも容易に入手可能となったことが挙げられる。また、別の要因として首都圏を中心に両替機、自動販売機などの機械処理装置を対象とした偽造券行使事件も多く見受けられた。その後、2006年の銀行券の改刷に伴い、偽造の発見枚数は減少したが、この事例は技術の発展によって新たな技術が生み出されることがきっかけとなることで、急激に偽造が増加する事象の発生は常にあり得ることを示している。

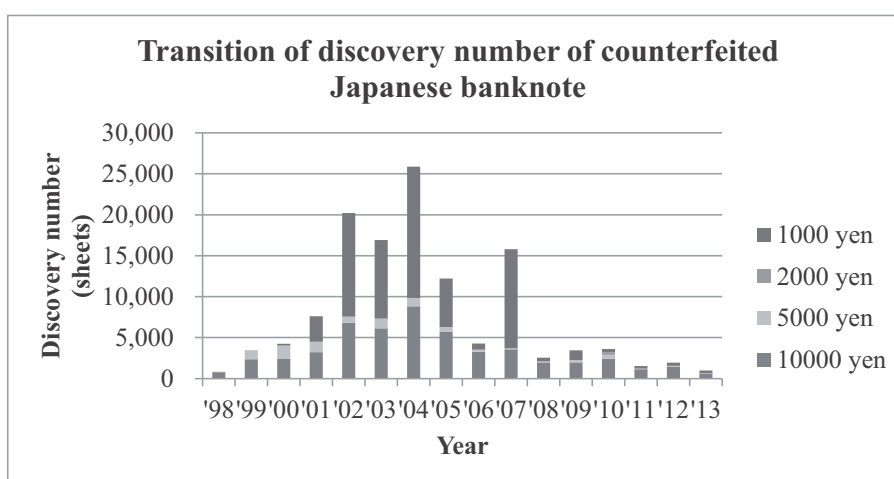


図1.2 偽造日本銀行券の発見枚数の推移 (出典： [5] [6], 警察庁)

Fig. 1.2 Transition of discovery number of counterfeited Japanese banknotes

(Ref.: National police agency)

1.4 紙と貴重印刷物

貴重印刷物には、その偽造を阻止するために各種の偽造防止技術 [7]が実装されている。偽造防止技術は、その真偽判別の方法によって3段階のレベルのグループに分けられる [8] [9]。各分類と各分類における偽造防止技術を以下に示す。

① 第1次認証 (First-line inspections)

特別な道具を用いることなく、一般人の官能に基づいて真偽判定が可能な技術。すき入れ¹、凹版印刷²、エンボス加工、表裏合成模様³、穿孔加工⁴、セキュリティスレッド⁵、ホログラム⁶、光学的変化インク⁷などがある。透かす、触る、傾けるなどの簡単な動作で確認可能である。例として日本銀行券に付与されている第1次認証技術の一部を図1.3に示す。

② 第2次認証 (Second-line inspections)

道具を用いて特殊な効果を発現させ、その効果を官能に基づいて真偽判定を行う技術。一定の訓練を受けた銀行の窓口業務者、および小売店の現金取扱者などによって認証を行う。微細画線、マイクロ文字、フォトクロミックインク⁸、サーモクロミックインク⁹、発光インク¹⁰、磁性インク、バーコードなどがある。人によるルーペ、紫外線ランプなどの簡単な道具の行使に付け加え、銀行ATMや銀行券両替機上に磁気測定プローブなどの形で実装される場合もある。第2次認証技術の例を図1.4に示す。

③ 第3次認証 (Third-line inspections)

高度な機器を用いて実行される真偽判定。製造元の研究所や鑑定施設の専門家によって判定される。特殊な技術であるため、その技術が実装されていることが非公開な場合が多い。主に使用される機器として、分光光度計、顕微鏡、赤外放射計、赤外線カメラ、特殊な化学指示薬の検出装置などがある。

¹ JISでは紙を透かして見たときに現れる模様と定義されている。

² 凹んだ画線部にインキをつめて強圧によって紙にインキを転写する印刷方式。盛り上がりのある画線部を手触りで識別できる。

³ 表裏同時印刷が可能な印刷機を用いて正確な刷りあわせで作製された模様のこと。透過光で観察すると表裏の模様が合成されて別の画像が観察できる。

⁴ レーザなどを用いて微細な穴を基材に開けること。

⁵ 製紙工程で思想に漉き込まれた偽造防止技術が付与された細かいフィルムのこと。

⁶ 生物体から出る光波と、それと干渉性のある光波との干渉パターンを記録材料に記録したもの。

⁷ 薄膜状粒子の光の干渉によって観察または照明角度の違いによって色が変わるインキのこと。

⁸ 太陽光や紫外線ランプによって色が変わる。

⁹ 指先による加熱などによって色が変わるインキのこと。

¹⁰ 紫外線などを当てると発光するインキのこと。

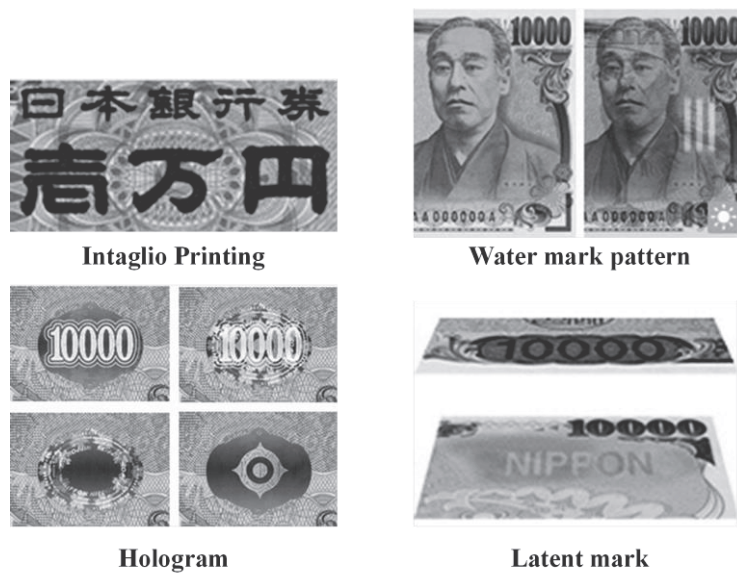


図 1.3 第 1 認証技術の例 (出典: [10], 国立印刷局)

Fig. 1.3 Examples of First-line inspection technology (Ref: National Printing Bureau)

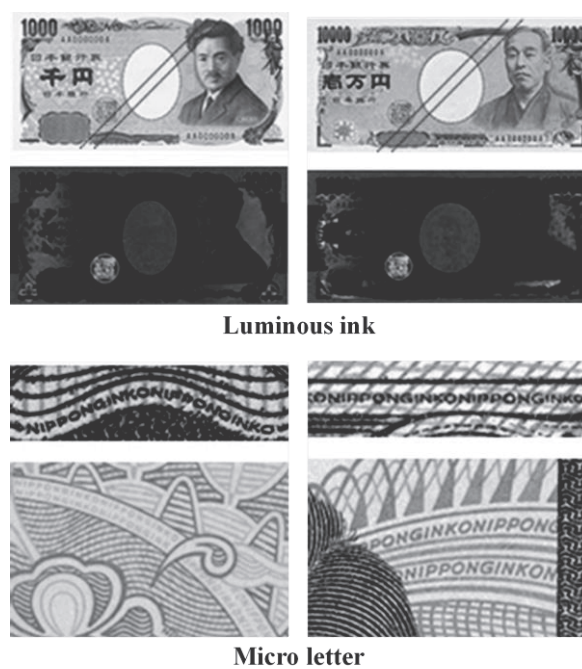


図 1.4 第 2 認証技術の例 (出典: [10], 国立印刷局)

Fig. 1.4 Examples of Second-line inspection technology (Ref: National Printing Bureau)

上記のどのレベルの検証技術においても基材のもつ特性は重要な役割を果たす。例えば、すき入れやエンボス、穿孔加工などの特殊加工においては、加工に合わせた適性が求められる。特殊なインキやホログラムの付与においては、安定した定着性が求められる。その他にも各偽造防止技術の効果の付与または発現において、基材の持つ特性は大きな影響を与える。このような貴重印刷物の基材には、紙媒体が多く使用されている。例えば、貴重印刷物の代表格である日本国における銀行券および旅券に使用されている基材は紙である。

代表的な偽造防止技術と紙の関係として銀行券に付与されているすき入れ（図 1.5）を例に挙げる。すき入れは偽造品と真正品を判定する際によく利用されている偽造防止技術とされ、また、偽造者によって再現が難しい技術として、日本国だけでなく多くの国の銀行券に採用されている。日本銀行券においては、白黒すき入れ¹¹によって豊かな濃淡階調とコントラストを持った人像模様が付与されている。このすき入れの付与において紙の特性はその品質に大きな影響を与える。影響の大きい紙の特性の一つである地合¹²が異なる紙の例を図 1.6 に示す。図 1.6 に示した画像は製造条件の異なる紙を透過型スキャナで撮像した画像である。画像上に見える濃淡のムラが地合であり、このムラの程度が小さい紙ほど地合の良い紙とされる。地合の変動は材料や製造条件の変動に起因して発生する。紙の地合は、紙を透かして確認した際にすき入れの背景部となるため、地合の悪化はすき入れの視認性を著しく低下させる。また、地合は繊維の凝集のムラでもあることから、紙の製造においてすき入れの付与時においてその品質へ悪影響を及ぼす。このため一般的な紙と比較すると、日本銀行券用紙では地合の管理はより重要とされている。

¹¹ 透かして観察した際に他の部分より白く見える部分を白すき入れ、黒く見える部分を黒すき入れと呼ぶ。日本銀行券に使用されている黒すきはすき入紙製造取締法によってその製造を制限されている。日本銀行券、公債証書、収入印紙、その他政府の発行する証券にすき入れてある文字もしくは画紋と同一若しくは類似の形態の文字若しくは画紋を白くすき入れた紙は、政府、独立行政法人国立印刷局又は政府の許可を受けた者以外の物は、これを製造してはならない。

¹² 紙を形成している繊維の分散の状態を表す用語。紙の繊維が均一に分散して、透かして観察した際に透過光量にムラがなく均一に見える紙が「地合の良い紙」、逆に不均一な紙を「地合の悪い紙」と呼ぶ。




	E 千円券 (野口英世)
	E 五千円券 (樋口一葉)
	E 一万円券 (福沢諭吉)

図 1.5 日本銀行券の白黒すき入れ

(出典： [10], 国立印刷局)

Fig. 1.5 White and Black water mark of Japanese banknote

(Ref: National Printing Bureau)

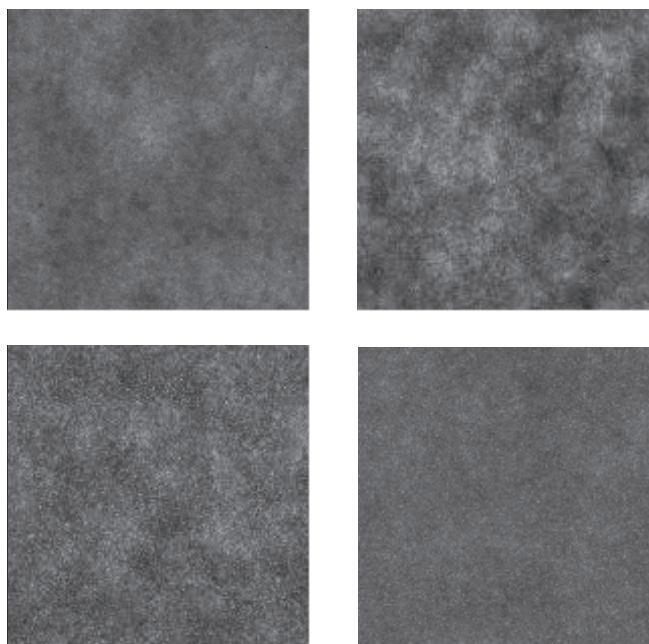


図 1.6 紙の地合の違い

Fig. 1.6 Difference of paper formation

1.5 貴重印刷物用紙の製造と品質管理の課題

貴重印刷物用紙の製造工程として、日本銀行券の製紙工程の概要 [11]を表 1.1 に示す。裁刻工程から調合工程までは完成紙料¹³を作成するための紙料工程であり、抄造工程は紙料工程で作製された完成紙料から紙を製造する工程である。

まず、裁刻工程では主体材料であるパルプシートを裁刻機でそうめん状に裁断する。一般的な紙の製造では裁刻処理は行われず、パルプシートはそのまま離解工程で処理されるが、日本銀行券用紙においては、精巧なすき入れの付与および地合の向上のため、パルプシートを切断し繊維の長さを調整する。次に離解工程において、離解機¹⁴によって繊維を 1 本 1 本ほぐし、水の中に分散させ、懸濁液状に加工する。その後、精選工程において異物を取り除く。精選工程ではリフラー¹⁵、スクリーン¹⁶、クリーナー¹⁷など複数の装置を組合せて様々な大きさの異物を除去する。この工程で異物を除去できなかった場合、紙料中に混入することになり異常製品の原因となる。次に叩解工程において繊維を加工する。紙の品質は叩解処理で決まると言われるほど、製紙工程における叩解は重要な処理であり、日本銀行券用紙の地合、強度、透明性、伸縮などの物性に大きく影響する。叩解処理では叩解機¹⁸によって繊維をフィブリル化¹⁹させる。叩解が進むにつれて紙は高密度化し、用紙強度は増加するが、引裂強度は低下する。また、極度に叩解を進めると抄造工程において、ワイヤー上での水切れが悪くなり支障をきたす。次に調合工程において、これまでの工程で処理された複数種類のパルプ、てん料²⁰、薬品（サイズ剤²¹、湿潤紙力増強剤²²、紙力増強剤²³など）を所定の割合で調合機に投入し混ぜ合わせる。この処理によって日本銀行券用紙の完全紙料が仕上がる。その特性は、ろ水時間や繊維長の分布などで管理される。最後に抄造工程で抄紙機によって日本銀行券用紙を製造する。抄紙機では、ワイヤー上で紙層

¹³ 製紙工程において紙を構成するための材料や薬品等が調合され、抄造を行うことが可能な最終状態となった紙料のこと。

¹⁴ 水を張った槽内においてローター刃などを回転させて回流を発生させながら繊維をほぐす装置。

¹⁵ 蛇行した経路の途中に堰板やマグネットを設置し、紙料中に含まれる砂や鉄片などの重い異物を沈殿分離させる装置。

¹⁶ スリットが設けられた平板状のプレートを上下振動させ、スリットを通過できない軽くて形の大きい異物を除去する装置。

¹⁷ 円錐状のパーツへ紙料を流し入れ、内部で渦巻き状の流れを作り遠心力を発生させることにより、繊維より比重が大きい微細な異物を除去する装置。ノズル径の大きなものから小さなものへ複数回通過させることで微細な異物を除去し、紙料のみを回収する。

¹⁸ 刃を持った 2 枚のディスクの間に紙料を通過させ、すり潰しながら繊維をほぐす装置。材料の種類や処理の程度に応じてディスクの枚数や刃の形状を変更する。

¹⁹ 繊維を毛羽立たせ、ささくれさせること。細毛化ともいう。一般的にはフィブリル化が進むほど繊維が絡みやすくなり紙の強度が増す。

²⁰ 繊維間の隙間を埋め、紙の不透明度、白色度、平滑度を向上させる添加物のこと。

²¹ 紙への水の吸収、浸透を防止または遅らせる。水性インキのにじみを防ぐ薬剤のこと。

²² 水分を含んだ状態の紙の強度を向上させる薬剤のこと。

²³ 乾燥状態の紙の強度を向上させる薬剤のこと。

を形成し、脱水部で水を絞り、その後ドライヤー部で乾燥し、用紙の強度や表面状態を調整するための薬液の塗布が行われ、最後に光沢を出す処理が行われた後、ロールへ巻き取られる。抄造工程における各部では紙の品質を目視や機械検査によって管理しながら製造は行われる。

紙料工程は、均質な紙を製造するための基盤となるため、目的に合った均一な紙料を安定して抄造工程に供給することが重要である。しかし、紙の工程は非常に多くの処理で構成されているため、各工程の変動が影響し合い、紙料を均一に保つことは容易ではない。また、抄造工程における抄紙機の状態、湿度、温度などの製造時の環境も紙の品質に影響を与えるため、さらに紙を形作る際の物理現象は複雑となる。実用的な対策として各工程において、規格化された評価方法などを用いて目視や測定機器によって特性を管理することが、その品質を管理する上で有効である。一般的な用途の紙においては、上記のような管理によって必要な精度の品質管理は可能であると考えられる。しかしながら、貴重印刷物の基材としての紙媒体には、高い偽造防止効果を維持し、偽造を防止するという使命がある。人の目に対して認識のし易い偽造防止効果の発現はもちろん、機械読み取りに対する安定性もますます重要とされてきている。近年では、国内における自動販売機の台数（図 1.7） [12]および自動窓口業務処理機（Auto Teller Machine; ATM）は市中において飽和するほどに普及し [13]、人から機械へ現金を支払うことは日常茶飯事となり、銀行での現金の取り扱いにおいてもその多くは ATM を通して行われている。また、食料品店などの小売店のレジにおいても、現金を自動的に判別および計測する自動釣銭機が多く見受けられるようになってきている。このような状況から、貴重印刷物の基材となる紙媒体の製造者には、紙の品質をより正確かつ詳細に位置づけ、その品質と製造工程における影響な関係を把握し、紙の特性をより高精度に管理する方法を確立することが求められている。

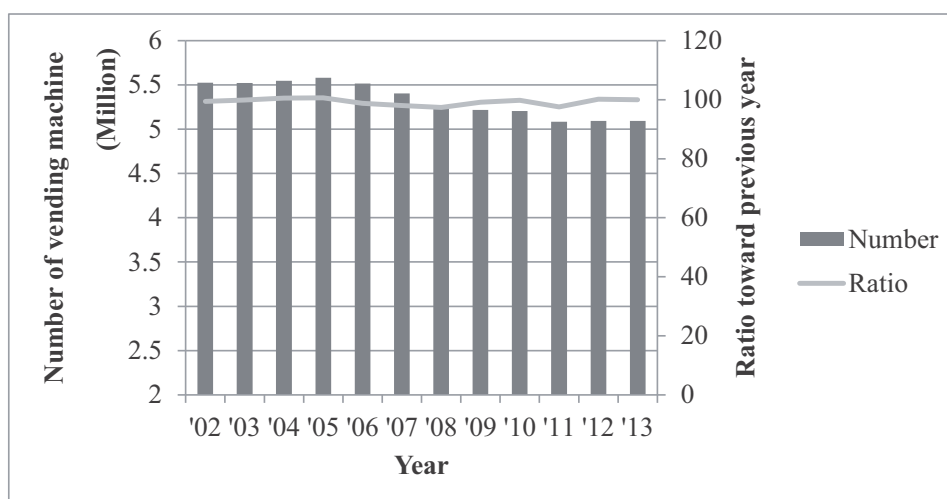








図 1.7 国内における自販機台数の推移（出典： [12]，日本自販機工業会）

Fig. 1.7 Transition of number of vending machine in Japan
(Ref: Japan vending machine industrial association)

表 1.1 日本銀行券の製紙工程（出典： [11], 国立印刷局）

Table 1.1 Production line of paper for Japanese banknote

(Ref.: National Printing Bureau)

1 裁刻工程		<p>紙の材料であるパルプシートを裁刻機（左写真）で細かく刻む。一般的な用紙の製造ではパルプシートに裁刻工程は行われませんが、日本銀行券用紙においては精巧なすき入れおよび良好な地合を得るために繊維を短くする。</p>
2 離解工程		<p>裁刻処理によって短くなったパルプを、水を張った離解機（左写真）に投入し、回流などによって繊維を1本1本に解きほぐす。</p>
3 精選工程		<p>複数の機械を組合せて様々な形態の遺物を取り除く。クリーナー（左写真）では円錐状のパーツにおいて遠心力を利用することにより、繊維より比重が大きい微細な異物を除去する。</p>
4 叩解工程		<p>叩解機（左写真）によって繊維をすりつぶすことによって繊維を毛羽立たせる（フィブリル化）。この処理によって繊維同士を絡みやすくする。</p>
5 調合工程		<p>紙の材料と薬品などを決められた配合割合で調合機へ投入し混ぜ合わせる。この作業によって抄造が可能な紙のもとである「完成紙料」となる。</p>
6 抄造工程		<p>完成紙料をワイヤーの上に流して薄い紙の層を作り、その後ドライヤー部で乾燥させて巻き取られる。すき入れはこの工程で付与される。</p>

1.6 研究の目的

貴重印刷物には、その基材として紙媒体が多く使用され、その紙には一般的な紙と比較して高い品質、そして特殊な特性が求められる。技術の進歩に伴い高精度な測定装置やその測定データから特徴量を算出する方法が開発され、従来と比較して紙の製造に関する多様な情報が取得可能となったが、紙の特性に影響する要因は複雑で、現在においても詳細な紙の特性および紙が構築されるメカニズムを正確に把握することは難しい課題である。

従来、このような課題においてはデータ間の相関係数の参照や主成分分析 (Principal Component Analysis; PCA)、重回帰分析 (Multiple Linear Regression; MLR) [14]などの統計手法が多く用いられる。このような統計的なアプローチは手法としてよく確立され、実用的な問題への適用が容易である。また、その解析結果もシンプルで、結果と要因の関係の説明が容易である。上記のような特徴から、統計的手法は従来から広い分野において使用されている。しかしながら、このような統計手法は複雑な問題への適用が難しいという問題がある。一方、近年、パーソナルコンピュータの進歩に伴い、進化計算や神経回路網 (Neural Network; NN)、サポートベクターマシン (Support Vector Machine; SVM)などを代表とした機械学習によって大規模データを解析する手法が活発に研究されている。機械学習はコンピュータの計算処理によって自動的に情報の探索 (学習) を行い、目的の処理を行う情報処理 (モデル) を構築する。学習を通じて自動的に大規模データの中から、人が気づくことが難しい基準やルールを発見し、複雑なモデルを柔軟に構築可能であるため、様々な問題に適用可能である。紙の製造に関する従来研究として、紙の特性の解析における実問題へ機械学習を適用した例もある。しかし、一般に機械学習は自由度の高いモデルを構築可能である反面、一方でその構造が複雑になり、結果と要因の関係の説明が困難となる。このため、実問題への適用において、機械学習によって作製されたモデルが示す結果の信頼性を保障することが難しいという問題がある。具体的には製品ラインにおける検査や貴重印刷物の真偽判別において機械学習によるモデルを使用すると、機械学習のモデルを使って得られた結果について、その結果に至った根拠や理由を求められる。特に、構築されたモデルがイレギュラーな結果を示した場合には、その原因の特定は強く求められる。このため、モデルの持つ結果の正確さと同時にその結果に至る過程の説明も重要とされ、製造工程で実際に運用する際には必要不可欠である。

そこで本論文では、紙の製造において取得可能となった多様かつ大量の情報を活用する方法として機械学習を用い、紙の特性をより高精度に分類、予測することが可能でかつ、人に理解し易い紙の分類および特性予測モデルを構築することを目指す。

第2章 紙と紙料の評価方法

2.1 製紙に関わる一般論

本節では、本論文の内容に関わる製紙の一般論および基礎知識について述べる。

2.1.1 紙の定義

日本工業規格 (Japan Industrial Standards; JIS) における JIS P 0001 [15]では、‘植物繊維その他の繊維をこう着させて製造したもの。なお広義には、素材として合成高分子物質を用いて製造した合成紙のほか、繊維状無機材料を配合した紙も含む。’と定義されている。つまり、植物などの繊維を抽出し、これを水中に分散させ、繊維を絡ませながら薄く平らに成形し乾燥させたものを“紙”という。より具体的に表現すると、紙は幅 20~40 μm 、長さ 1~3mm、厚さ数 μm からなる主として木材パルプ繊維が 10 本から 100 本以上絡み合い、その一部は結合状態を形成してシート状を構成し、その構成中に薬品や場合によっては数 μm のてん料、異種の素材が加わって出来上がっている薄層体である [16]。したがって、紙の本質を理解するためには、紙を構成している単繊維とその他の添加材料が集成して形成しているマトリックスの相互作用とそれらの関係を理解することが必要である。

2.1.2 紙の分類

紙はその材料や製造方法によって、洋紙、板紙、和紙、合成紙などに分類され、さらに用途によって印刷用紙、筆記用紙、ろ紙、電気絶用紙など細かく分類される [17]。また、紙の pH によって、酸性紙と中性紙に分類される [18]。紙の pH は紙の製造工程で加えられる薬品の違いによって変化する。1850 年代以降の製紙工程で、インキのにじみ止めの定着剤として硫酸アルミニウムが使用された紙を酸性紙と呼ぶ。この硫酸アルミニウムが加水分解され、硫酸を生じて紙の繊維を傷め、紙を劣化させる。1000 年以上たった書物が残っている中、比較的近年に作成された書物のうち、変色を伴いながら図 2.1 のように著しく劣化している書物には酸性紙が使用されているものが多い [19]。以上のように、紙の用途やその特性は極めて広範囲にわたっており、一言で紙と言っても、その用途に合わせて材料や薬品などを選択し、適切な製造方法で製造することによって、全く異なる特性を持った基材として活用することができる。



図 2.1 酸性紙の劣化 (出典： [19], 東京都立図書館)

Fig. 2.1 Degradation of acid paper (Ref: Tokyo metropolitan library)

2.1.3 紙の材料

紙の主な原料は木材であるが，一部では非木材も用いられている．木材では広葉樹パルプ (Leaf Bleached Kraft Pulp; LBKP または Hard wood Bleached Kraft Pulp; HBKP と呼ばれる) と針葉樹パルプ (Needle Bleached Kraft Pulp; NBKP または Soft wood Bleached Kraft Pulp; SBKP よばれる) が主な分類として挙げられ，広葉樹には，ブナ，ナラ，カバ，ポプラなどが挙げられ，繊維が短い傾向がある．針葉樹には，松，杉，ヒノキなどが挙げられ，繊維が長い傾向があるため，針葉樹を原料とした紙の強度は高いとされている．また，非木材の原料では，靱皮繊維とその他の繊維に分類され，みつまた，こうぞ，木綿，ケナフ，アバカなどが挙げられる．日本銀行券の原料としては，アバカとみつまた (図 2.2) が使用されている [20]．みつまたは古くから和紙の原料として使われており，1879 年に日本銀行券の用紙の原料として採用されてから現在まで使用されている．海外の銀行券においては，綿などを使用する国が多く [21]，日本国銀行券とは原料が異なる．



図 2.2 日本銀行券の原料 (出典： [20], 国立印刷局)

Fig. 2.2 Raw materials of Japanese banknote (Ref: National Printing Bureau)

2.1.4 紙の基本的性質

紙の持つ基本的な性質には、水または湿度の変化に対する性質である親水性、油の吸収における性質である親油性、繊維の配向によるカールのし易い方向を示す方向性、表面の平滑さである平面性、光または気体などに対する透過性、圧力を加えた際に形状の復元の度合いを表す性質である圧縮性、そして耐久性などが挙げられる。上記のような紙の性質は、繊維の密度および空隙（図 2.3）に大きく支配される [16]。紙は、繊維、空隙、てん料、その他の添加物で構成されており、その中の空隙、てん料、添加物の占める割合によって紙の性質は大きく変わる。空隙の大小、分布によって紙の通気性、不透明性、インキ、ワックス等の吸収性、表面の構造が変化する。

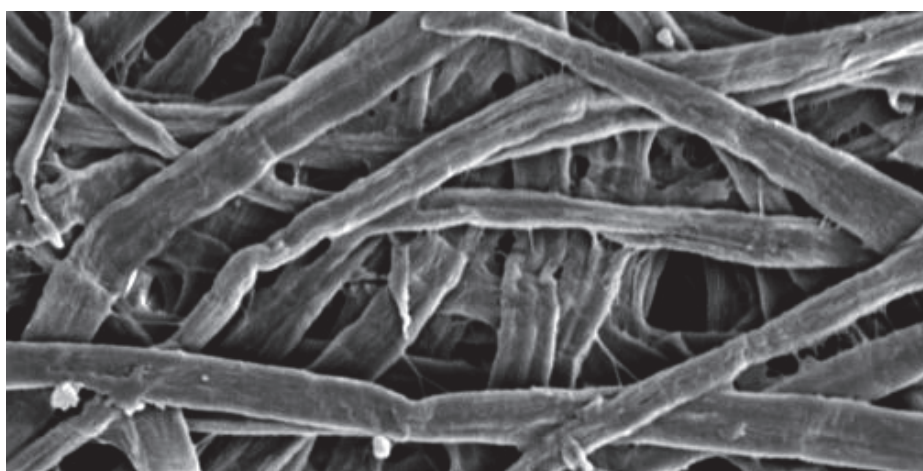


図 2.3 紙における繊維の状態（電子顕微鏡を用いて撮像）

Fig. 2.3 Appearance of Fibers in paper (captured with Scanning Electron Microscope)

2.1.5 紙の特性を表す指標

紙の製造を行う上で、その特性を管理するための様々な指標がある [22]。紙の特性を表す指標は多岐にわたっており、紙の用途に応じて管理する指標やその値が異なる。本項では紙の特性を表す指標のうち、代表的なものについて述べる。

① 水分

紙に含まれる水分は紙の質量、強度、寸法安定性、耐久性、電気的性質などに影響し、印刷適性においても重要な因子である。

② pH

紙の pH は耐久性、とりわけ保存性に非常に大きく影響する（2.1.2 項参照）。また、印刷インキの乾燥速度にも影響することから加工原紙において重要な指標である。

- ③ 灰分
紙を燃焼して得られる灰の質量。灰は主として填料または顔料で構成されていることから、それらの量の指標として使用される。
- ④ 坪量
紙の質量を表す一般的な指標で、単位面積当たりの重さである。通常 1m^2 当たりのグラム数をその紙の坪量 (g/m^2) とする。
- ⑤ 厚さ
規定された一定の圧力下においたときの紙の厚さ。坪量と合わせることによって紙の密度を表す指標としても用いられる。
- ⑥ 引張、引裂、耐折、破裂強さ
引張強さは張力に対する強さ、引裂き強さは引裂きに対する抵抗性、耐折強さは紙の耐えうる折り畳みの回数、破裂強さは圧力に対する強さを表す。耐久性を評価する上で重要な指標である。市中を還流する銀行券用紙においては通常の印刷用紙と比較して高い強度が求められ、特に折りたたまれることが多いことから高い耐折強度が付与される。
- ⑦ こわさ
紙の曲げに対する抵抗であり、“コシ”ともいわれる。印刷用紙においては取扱い上適切なこわさが必要とされる。日本銀行券は通常の印刷用紙と異なるこわさを持っており、人の知覚における真贋判定の際に一つの指標として用いられることもある。
- ⑧ 平滑度
紙の表面状態に関する指標の一つであり、インキの転移や定着に影響するため、印刷適性においての重要因子の一つである。
- ⑨ 白色度
紙の白さに関する指標で、紙を不透明になるまで重ねて青色光を照射した際の反射光の強度で表される。
- ⑩ 不透明度
印刷用紙において重要な性質であり、紙に白色体を裏当てたときの反射率と黒色体を裏宛した時の比で表される。

⑪ 地合

紙中の繊維の分布の均一性を示す値で、紙を光に透かして見たときの繊維の分布状態の均一性で評価される。

上記のように紙の特性は大きく、構造的な性質、物理的な強度に関する性質および光学的性質に分けられる。坪量や灰分などの構造的な性質は紙の詳細な組成を把握する指標、引張強さやこわさ物理的な強度に関する性質は耐久性に関わる指標、白色度や地合などの光学的性質は見た目や質感に関わる指標として主に用いられる。各特性を測定するための試験方法および評価方法については2.2節で述べる。

2.1.6 製紙用パルプの特性を表す指標

紙の特性と同様、紙の材料である製紙用パルプにおいても用途に合わせて異なる指標が用いられる。ここでは製紙用パルプの特性を表す基本的な指標 [22]を記載する。

① 繊維長

製紙用パルプの特性を示す基本的な指標である。パルプにおける繊維長は幅広い分布を持っており、平均繊維長は数平均繊維長 $L(n)$ 、長さ加重平均繊維長 $L(length)$ 、質量加重平均繊維長 $L(weight)$ のいずれかで比較される。一般にシートの強度は繊維長が長い程強く、短い程弱くなる。紙の強度への影響は一般に長さ加重平均繊維長や重さ加重平均繊維長を用いて説明される。以下に各指標の算出式を示す。

$$L(n) = \frac{\sum_i l_i}{n} \quad (2.1)$$

$$L(length) = \frac{\sum_i l_i^2}{\sum_i l_i} \quad (2.2)$$

$$L(weight) = \frac{\sum_i l_i^3}{\sum_i l_i^2} \quad (2.3)$$

② 繊維の粗度

単位長さ当たりの繊維の重量で表される。紙の強度、液体の浸透性、表面粗さ、印刷適性に影響のある指標である。

③ ろ水度（ろ水時間）

パルプにおける水切れを表す指標であり、一定量のパルプの水が切れる時間などで表される。叩解の度合いの指標として用いられる。抄造装置のワイヤー上での水切れ状態は、均一な紙の製造において重要である。

上記のように製紙用パルプの基本的な特性としては繊維に関わるものが多くあげられ、紙の特性を決定づける上で重要である [23]. 上記以外にもパルプ繊維における水分や灰分なども指標として用いられる.

2.2 紙の評価技術およびその研究

紙の特性を評価する一般的な試験方法については JIS, 国際標準化機構 (International Organization for Standardization; ISO), パルプ製紙業界技術協会 (Technical Association of the Pulp and Paper Industry; TAPPI) および JAPAN TAPPI 紙パルプ試験方法などにおいて, 規格化 [24] [25]されている。それぞれ専用の方法や装置を用いて紙の構造や物理的特性を測定するものである。代表的な試験方法について 2.2.1 項で述べる。また, 紙の特性をより詳細に評価するため, あるいは製造ライン上で紙の品質を管理するために, 上記のように規格化された試験法以外の手法も研究されている。紙の特性の測定方法については, さまざまな研究がおこなわれているが, その中に, 紙やその材料を光学的に測定することによってその特性を評価する方法がある。光学的な評価は, 貴重印刷物における紙の識別性または機械読取性を評価する上で重要である。一般的に用いられる測定方法としては, 光源を紙で遮った際に発生する透過光を観察する方法が多く用いられる。透過光による評価では, 紙の地合や配向, 厚みや密度などが評価される。旧来の装置や, 速度が求められる製造工程上のオンラインの検査装置では, 装置の性能的制約のため, ポイントセンサなどによって部分的に光学特性を測定することもあるが, 近年では技術の進歩に伴ってラインセンサ, エリアセンサ, コンタクトイメージセンサ (Contact Image Sensor: CIS) を用いることによって広範囲の画像を撮像することが可能な装置が広く普及している。このため, 紙の光学特性の評価においては, 紙の画像を用いた評価を行うための環境は整っていると見える。以上の理由から本研究では光学的な方法によって得られる紙の特性を用いることとした (第 4 章, 第 5 章, 第 6 章)。

2.2.1 紙の評価方法

紙の特性に関する各種指標を測定する試験方法のうち, 基本的な規格を記載する。以下は JIS によって規格化されている試験方法である。

- ① 試料の前処理および試験条件(JIS P 8111 [26])
- ② 坪量の測定 (JIS P 8124 [27])
- ③ 厚さおよび密度の測定 (JIS P 8118 [28])
- ④ 紙の pH の測定 (JIS P 8133-1 [29], JIS P 8133-2 [30])
- ⑤ 水分の測定 (JIS P 8127 [31])
- ⑥ 灰分の測定 (JIS P 8251 [32])
- ⑦ 引張強さの測定 (JIS P 8113 [33])
- ⑧ 耐折強さの測定 (JIS P 8115 [34])

上記以外の評価方法として、地合いの評価方法がある。地合いの評価は銀行券などのすき入れが付与される貴重印刷物においては重要な指標である。その評価方法としては主観的な方法と客観的な方法の両方が現在でも用いられている。主観的な方法では、ライトテーブルなどを用い、紙を光に透かしてその濃淡ムラを視覚的に判断し評価する。客観的な評価方法としては、光源とスキャナやラインカメラまたはエリアカメラなどの撮像素子を用い、紙の透過光を走査することによって画像を撮像し解析する方法や β 線ラジオグラフィによって軟 X 線フィルムに微小な坪量変動を濃度変化で表し、その濃度変化を受光素子で読取り画像を形成し、解析する方法 [35]が提案されている。スキャナ等を用いて紙の透過光を撮像方法は紙の色や不透明度の影響がノイズとなる。一方、 β 線を用いた方法質量以外の性質には依存しないため、より正確に地合いを評価することが可能である。しかしながら、 β 線を用いた測定は容易ではなく、実用的にはカメラ等を用いて紙の透過光を測定する方法が多く用いられる。カメラ等を用いる方法では、撮像した画像に対して画像処理を行い、紙の光学的特性の指標として用いる。 β 線等を線源とする質量地合いと呼び、光を透過させたときに観察される地合いを光学地合いと言う。紙を透過する光量は紙の密度により大きく影響されるため、カレンダー処理²⁴などを行った紙では、光学地合いと繊維の分散状態は一致しないことがある。近年、光学地合いを測定する装置が多く開発されているが、測定方式や特徴量の算出方法において統一的な方法はない。このような現状に合わせて、製造工程における熟練作業者の視覚的な地合いの評価と各種測定値および特徴量との相関が問題となることも多い。

2.2.2 紙の評価に関する従来研究

前項のように規格化された試験方法がある一方、製造工程においてより精度の高い評価や実用的な評価を行うため、紙の評価方法に関する様々な研究が行われている。以下に本研究と関連のある従来研究を記載する。

① コンピュータ画像処理によるパルプシートの構造解析

撮像した画像からの特徴量算出には、スペクトル解析や同時生起行列 (Co-occurrence Matrix; CM) が人による評価との相関が高く有効であるとされている。湯原らによって上記のような画像から算出可能な特徴量を用いた紙の評価手法の有効性が示されている [36] [37] [38]。

② 紙の強度特性の予測に用いる超音波測定の開発

紙の強度を評価するための超音波を用いた評価手法がある。Allan らの研究によって従来の試験方法より簡便に紙の強度を表す指標を算出可能である手法が提案されている

²⁴ 紙の表面に金属のローラーで圧力をかけ光沢や平滑性を高める処理のこと。

[39].

③ 紙の内部散乱光特性の評価

滝代らによって、紙の内部散乱特性を紙の光学伝達関数（Modulation Transfer Function ; MTF）として評価する測定装置が開発されている [40]. 紙の内部散乱光特性は、印刷における光学的ドットゲイン現象を説明できる物性値であることが知られている. 提案装置はサイン波像のテストパターンを投影する投影装置と投影されたサイン波像のテストパターンの反射光強度分布を計測する CCD カメラからなる. 実験では、MTF による評価によって光内部散乱を良く説明することができる結果が得られている.

④ 紙表面の印刷適性の数値評価

画像処理によって紙の密度を測定し、紙の品質を数値的に予測する方法が Danby らによって提案されている [41]. 提案手法では紙における白紙品質だけでなく印刷品質をも予測することが可能であることが報告されている.

⑤ 紙の高分解能測定からの紙の光学的変動の予測

坪量、密度およびてん量の高分解能測定による値から、紙の光学的特性を予測する研究が Hägglund らによって行われている [42]. 提案手法における予測モデルが示す結果は、視覚的な反応と同様の結果が得ることが可能であり、また、てん量の影響と比較して坪量のほうが紙の光学的特性への影響が大きかったことが報告されている.

⑥ 紙の光学特性の品質評価における画像解析

紙を低角度の光源照射を用いて撮像することによって得られる画像から紙の光学的品質を評価する手法が提案されている [43]. 提案手法を用いた実験では、入射角 17.5 度と高分解能の画像の組合せから、紙の白色度などのパラメータを良好な精度で予測可能であったことが報告されている.

⑦ 紙の地合の測定手法に関する研究

高分解能の光学スキャナを用いて撮像した画像において、画素値の平均値と分散値から算出可能な変動係数と画素値の明るさの範囲である SP (Specific Perimeter) という指標を作成し、2つの値を2次元プロットしたマップを地合マップとして、地合を評価する方法が Dilleby によって検討され、銘柄の違いやコート紙の表面特性の違いの区別において有効性が示されている [44].

2.3 製紙用パルプの評価技術およびその研究

紙の試験方法と同様に JIS や ISO などにおいて、紙パルプの試験法として規格化されており、それぞれ専用の装置を用いて化学的特性や物理的特性を測定する。その中でも特に重要なのは繊維の特性である。従来は、繊維の特性は主にろ水度の測定で行われていたが、繊維の詳細な状態の解析は熟練を要す方法が多く難しかった。そこで、パルプの繊維一本一本の特性を解析する手法に関する研究 [45]が行われた。近年では、繊維の特性を測定する装置として繊維特性測定器 [46]などが開発され、材料中の繊維一本一本の画像を高速に撮像し、繊維に関する様々な特徴量を算出することが可能となり、従来の方法では表現が難しかった効果などの指標として用いられている。また、上記のような装置は実験室における手動操作を中心としたオフラインでの使用が多いが、近年では精度および再現性の向上に加え、自動化によるオンラインでの評価への対応が求められている [47]。本研究においても繊維の特性と紙の特性の関係について着目して実験を行った（第 6 章）。

2.3.1 繊維特性の測定方法

本項では繊維の特性の測定方法として用いられている主な方法のうち、本研究に関連の深い方法について述べる。繊維の特性は叩解処理に強く影響を受けて変動するため、製紙工程における叩解処理を調節する際の指標であり、紙の特性を決定づける重要な要素である。主に紙料におけるろ水度から推定する方法、繊維をふるい分けした上で各繊維をガラスライド上で拡大投影して計測する方法、光学的な測定装置を用いることによって繊維一本一本を撮像して解析する方法がある。

① ろ水度（JIS P 8121-1 [48], JIS P 8121-2 [49]）

ろ水度の測定方法にはシートマシンによるろ水時間の計測、連続パルプろ水度測定器による測定、ショッパーろ水度計による測定、カナダ標準ろ水度試験など、いくつかの方法がある [22]が、ここではショッパーろ水度計について述べる。試験器は底に目の開き 0.175～0.147mm の金網と内部に円錐座のあるろ水等および側管の付いた計測漏斗で構成される。測定に当たっては濃度を 0.2%に調整し、離解した紙料を 1000ml 用い、紙料を流下させたときの側管からの排水量を読み取る。類似した方法としてカナダ標準ろ水度試験があげられ、ショッパーろ水度計と同様紙料を流下したときの排水量を測定する。大きな違いは円筒部の底に金網ではなく、0.51mm の固定径の穴が多数配置されているところである。ショッパーろ水度は叩解の度合い、カナダ標準ろ水度計はろ水の度合いを測定するのに適している。

② ふるい分け試験法 (JIS P 8207)

複数のふるい分け層を持つふるい分け試験機 (TAPPI T 233) を用い、4 個のふるい分け層に目開きの異なる金網を目の大きいものから目の細かいものへ順にセットし、規定の濃度に調整したパルプ懸濁液を一定の条件下で流すことによって繊維をふるい分ける。各層にふるい分けられた繊維の量と、ふるい分けされずに排出された繊維の量の比をふるい分け残量として紙料中における繊維の特性の指標として用いる。また、より詳細な繊維の特性を解析する際には、5 つ以上の層で繊維をふるい分けし、その後、各層でふるい分けされた繊維を採取し、ガラススライド上でその長さや形状を測定する投影法 (TAPPI T 232 [50]) もある。

③ 光学的計測法 (JIS P 8226 [51], JIS P 8226-2 [52], ISO16065-1 [53])

光学的計測法では、繊維の偏光特性と CCD カメラを用いて懸濁液中に分散した繊維一本一本の画像を撮像し、その画像より繊維の形態を自動で解析・測定する手法である。繊維の撮像に置いては、極細の透明の管の中に繊維の懸濁液を通し、偏光を照射する。このとき繊維を通過した偏光は光軸が回転することから偏光フィルターを用いることによって繊維の画像のみを CCD カメラで撮像可能となる。繊維の懸濁液は連続的に容器から透明の管内へ送り出され、レーザーセンサなどによって繊維を検出することによって大量の繊維の画像を自動的に撮像し、その画像から繊維の長さ、幅、壁圧などを特徴量として算出する。また、非偏向な光源を使用して撮像する手法もある。

ろ水度やふるい分け残量は繊維の特性を表す重要な指標として従来から採用されているが詳細な繊維の状態を把握することは難しい。例えば、ろ水度は叩解の度合いを把握するための指標として重要であるが、その値が増える要因として、短繊維や微細繊維の増加、繊維の膨潤、フィブリル化など複数の要因が考えられる。このため、ろ水度だけでは繊維の具体的な変化や叩解によって繊維に与えられた効果を詳細に説明することができない。また、ふるい分け残量は繊維のおおよその状態を把握でき、叩解処理などにおける繊維の変化を把握する上で重要な指標である。しかし、ふるい分け残量には、繊維長の変化、繊維の膨潤、繊維のねじれ、カール、外部フィブリル化など複数の要素が影響する。このため、繊維の詳細な変化を把握することはできない。以上の点において、近年開発された光学的計測法は、繊維の特性をより詳細に把握することが可能である。そこで本研究では光学的計測法によって測定した繊維の特性を用いることとした。

2.3.2 紙料の評価技術に関する従来研究

製造工程において、より正確に紙料の特性を把握することによって、紙の品質や特性を安定させるため様々な研究が行われている。近年では、繊維の詳細な特性を解析する手法や、製造工程上においてオンラインで測定可能な手法の研究が多く行われている。以下に

本研究と関連のある従来研究を記載する.

① 繊維の特性の評価方法に関する研究

画像処理技術によるパルプの品質の制御と製品品質の予測に関する研究が Ruonala らによって行われている [54]. この研究では, 近年の高解像度デジタル撮像や画像処理ソフトウェアの画像技術の発展によって改良され新たな繊維測定方法を用いることによって, 繊維のフィブリル化の様子を自動的に測定する方法を検討している. 従来, フィブリル化に関する分析は実験室レベルでは可能であったが, 製造工程上でオンラインに分析する測定方法は確立されていない. そこで, 高解像度の光学システムから算出された繊維の複数の画像特徴量と主成分分析 (Principal Component Analysis; PCA) を用いて, 製造工程から採取した紙量から紙の特性を予測するモデルを構築する実験が行われている. 実験結果ではフィブリル化指数などの詳細な繊維特性を画像から算出した特徴量を用いることによってオンラインで測定する方法が提案され, その可能性が示されている.

② 光と音響技術の組み合わせを用いたパルプ懸濁液中の繊維形態の推定

パルプ懸濁液は繊維と微細繊維 (200 μm 以下の繊維) で構成されており, 懸濁液中の微細繊維の比率は紙の特性に大きな影響を与えるが, 研究室レベルの環境以外測定が難しいことから製造工程上でオンラインに測定を行うための適した手法について研究が行われている. Niemi らの研究 [55]では光と超音波測定を組み合わせ, 部分最小二乗 (Partial Least Squares; PLS) を用いることによって懸濁液中の微細繊維と繊維質量の比率を同時推定する装置が提案され, 優れた予測性能が得られたことが報告されている.

③ 繊維分布特性に基づく手すきシートの引張強さおよび密度の予測のためのモデル

パルプの繊維特性に基づいて紙特性を予測するためのモデルを構築する研究が Pulkkinen らによって報告されている [56]. この研究においては, ユーカリを原料としたパルプ紙料における繊維特性として, 繊維長, 繊維幅, 繊維壁厚み, カール, 粗度, 断面積, 容積および保水値を測定し, その後, 各パルプ紙料から作製した手すきシートの密度, 物理的強度および光散乱性などとの関係を解析し, 一定の成果が得られたことが報告されている.

④ 木材繊維特性と中密度繊維紙の関係のモデル化

SHI らによって, 曲げ強さ, 弾性係数, 内部結合強さを応答変数とし, 密度, pH などの木材性質および繊維長, 繊維粗度, 粒度分布などを予測変数として多変量線形回帰分析を行った結果が報告されている [57]. 報告では微細繊維の曲げ強さへの影響や長繊維と pH による弾性係数への影響, 繊維幅の吸水量への影響など繊維特性と紙の特性の関

連の深さが示されている。

⑤ 紙品質のリアルタイムでの最適化

Loberg らによって繊維特性として重要な繊維長とろ水度を製造工程において常に同様となるようにリアルタイムで最適化する方法について検証がなされている [58]. 統計的モデルを作製し、繊維特性をオンラインで調整することによって紙の特性を向上させることが可能であったことが報告されている。

⑥ 繊維壁厚みの影響

Braaten によってパルプ紙料における繊維の繊維壁厚みが、引張強さおよび光散乱性に強い影響を及ぼすことが示されている [59].

⑦ 紙繊維の末端の画像ベースの自動検出




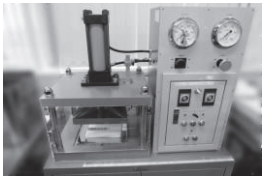


繊維の測定方法に関する研究において、個々の繊維レベルでの測定データを取得するための研究が Hirvonen らによって行われている [60]. 報告では、撮像した繊維の画像から画像処理技術によって繊維の端を正確に検出する方法が提案されている。

2.4 評価用手すきシートの作製方法

本節では、本研究において使用する評価用シートの作製方法について述べる。通常、紙の製造ラインは非常に大がかりで、少量のサンプルのみを作製することはできない。そこで小規模な実験の場合、一般に実験用の評価用紙は小型の装置を用いて手作業によって作製される（手作業によって作製された評価用紙を手すきシートと呼ぶ）。手すきシートの作製に当たっては、JIS（JIS P 8222 [61]）または ISO（ISO 5269-1 [62]）における規格がある。本研究では、実験用シートの作製は JIS および ISO の規格に則って実施した。表 2.1 に本研究における手すきシートの作製の手順および機器の一例を示す。主な処理としては、原材料の離解を小型の離解機、叩解を PFI ミルで行うことによってパルプ懸濁液を作成する。次に、パルプ懸濁液へ条件に応じて薬品等の投入や攪拌を行い、円筒型の手すきシートマシンによって紙層を形成する。その後、形成された紙層を圧搾装置によって圧搾、乾燥装置および自然乾燥を通して乾燥することによって手すきシートを作製する。

表 2.1 実験用手すきシートの作製手順の一例

Table 2.1 Example of procedure of creating hand-sheet

<p>1 離解</p>		<p>実験用の小型の離解機（左写真）を用いて、パルプシートなどをローターの回転運動によって離解する。この処理によってパルプシートなどを懸濁液状の紙料に加工する。</p>
<p>2 叩解</p>		<p>離解された紙料を小型の叩解機（PFI ミル）（左写真）へ投入し、叩解処理を行う。繊維のフィブリル化など加工の程度は、主に叩解回数の設定によって調整する。</p>
<p>3 シート形成</p>		<p>円筒型の手すきシートマシン（左写真）を用いて紙を作製する。円筒下部に円形のワイヤーを設置する、弁の開閉によって水が下部へ排出される仕組みとなっている。弁を閉め、円筒中に規定量の水を貯め、その中へ紙料を投入し攪拌する。その後、弁を開放することによって水が排出され、ワイヤー上に繊維の層が形成される。</p>
<p>4 圧搾</p>		<p>圧搾機（左写真）によって手すきシートから水分を取り除く、水を吸収する間紙の間に形成したシートを挟み込み、一定の圧力をかける。</p>
<p>5 乾燥</p>		<p>圧搾したシートを熱によってさらに乾燥させる。水分を吸収させるための生地と加熱したドラムに挟み込むことによって水分を除去する。</p>
<p>6 自然乾燥</p>		<p>最後に紙を自然乾燥させる。繊維の伸縮によって紙が歪むことを防止するためリングによって挟み込み、乾燥させ、その後、調湿された室内でシーズニングを行う。</p>

第3章 本研究に関する機械学習手法

3.1 機械学習 (Machine Learning)

機械学習は、機械（計算機）がデータの集合の中における既知の情報（教師データ）を基に、規則や判断基準などを自動的に獲得（学習）し、未知の情報へ対応する技術であり、代表的な手法としては進化計算や神経回路網などが挙げられる。機械学習には大きく分けて教師有り学習と教師無し学習の2種類がある。教師有り学習は、与えられた教師データを正しく出力することができる構造を獲得する手法であり、教師無し学習は、何らかの指標に対して適した割り当てを出力する構造を獲得する手法である。本研究では教師有り学習を主に用いている。

3.2 進化計算 (Evolutionary Computation; EC)

進化計算 (Evolutionary Computation; EC) は生物の進化から着想を得た最適化・探索手法であり、多点探索法の一つである。生物の進化プロセスである親個体から子個体へ遺伝情報の伝達を模倣し、遺伝子の組み換えや突然変異を探索点のオペレータとして用いることによって効率的に最適解または実用解を探索する手法である。進化計算法としては、Rechenberg らによって提案された進化戦略 (Evolution Strategy; ES) [63] [64], Fogel らによって提案された進化的プログラミング (Evolutionary Programming; EP), Holland らによって提案された遺伝的アルゴリズム (Genetic Algorithm; GA) [65] [66] [67], Koza らによって提案された遺伝的プログラミング (Genetic Programming; GP) [68] [69]がある (図 3.1)。本節では、本研究と特に関連の深い ES, GA そして GA の拡張である GP について述べる。

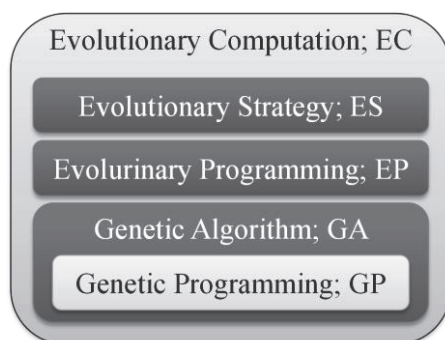


図 3.1 進化的計算法のグループ

Fig. 3.1 Group of Evolutionary Computation

3.2.1 進化戦略 (Evolutionary Strategy; ES)

ES は n 次元実数値空間における数値最適化問題に対する手法である。ES における探索点 (個体) は実数値ベクトル $\vec{x} \in R^n$ であり, \vec{x} を変数とする関数 $f(\vec{x})$ の最大値 (または最小値) を与える \vec{x} を求めることが最適化問題にあたる。ES では正規分布に従う乱数によって次の探索点を決定することによって探索を行う突然変異 (Mutation) を主なオペレータとする。各個体は目的関数によって得られる評価値 (適応度) を基に進化を続ける。進化では, 突然変異によって現在の探索点 S_1 の周辺に新たな探索点 S_2 を作成し, 目的関数によって適応度を求める。この際に S_2 の方が優れていれば S_2 が次の探索点となる。この処理を繰り返すことによって最適化処理を進める。ES においては μ 個の探索点 (親) から λ ($\geq \mu$) 個の探索点を生成して探索を行うが, その探索点の生成方法および決定方法にはいくつかの種類がある。代表的な手法としては, 生成した λ 個の子個体から選択する (μ, λ)-ES, μ 個の親と生成した λ 個の子を合わせた ($\mu + \lambda$) 個の個体から選ぶ ($\mu + \lambda$)-ES がある。その他の方法としてとして ($1 + \lambda$)-ES, ($1, \lambda$)-ES, ($\mu + 1$)-ES などがある。

3.2.2 遺伝的アルゴリズム (Genetic Algorithm; GA)

GA は J. H. Holland によって提案された生物の遺伝と進化の過程を模倣した多点探索に基づく探索アルゴリズムであり, 生物における淘汰, 交叉, 突然変異などの操作を用いて新しい個体 (探索点) を生成し, 実用解あるいは最適解を高速に発見する方法である。GA は現在までに多様な最適化・探索の問題に適用が可能であることから多くの実問題に適用され, その有用性が示されている。以下に GA における基本的な遺伝子操作, そして図 3.2 に各処理のフローを示す。

① 選択 (Selection)

選択は個体の適応度 (評価値) に基づいて, 次世代に残す個体を選ぶ操作である。選択手法としては, ルーレット選択やトーナメント選択が挙げられる。ルーレット選択では, 適応度に比例した確率で個体を選択する。これは適応度が大きい個体のみが次世代に残り, 個体群における多様性が失われ, 局所解にとらわれ易くなることを防止するための方法である。トーナメント選択ではランダムに選択した複数個体の中で最も適応度が高い個体を選択する。また, 上記の選択方法を拡張したエリート保存戦略がある。エリート保存戦略では各個体の適応度を基にソートを行い, 予め設定したルールに基づいて適応度の高い個体を優先的に生存させる方法である。いずれの方法においても, 適応度の高い個体を優先的に残し, 適応度の低い個体が淘汰される。

② 交叉 (Crossover)

交叉は実際の生物の生殖においてみられる現象を模倣した処理で, 選択された親個体

(現在の探索点)の組において、染色体の一部を交換して子個体(新たな探索点)を生成する操作である。交叉処理には、親に当たる個体の形質を子が継承する効果を期待している。交叉の方法には、染色体を1点で分割して交換する一点交叉や、2つ以上の点で分割して交換する多点交叉、各遺伝子をランダムに交換する一様交叉などがある。

③ 突然変異 (Mutation)

突然変異は染色体の内容を突然変異率 (Mutation rate) に従う確率でランダムに変更する操作である。突然変異に操作によって、交叉では生じない遺伝子をもつ個体を生成することができる。探索の観点からみれば現在の探索点から大きく離れた場所に探索点を発生させることができる。突然変異には局所解にとらわれた際に探索点をそこから脱出させる効果を期待している。交叉による親の形質を維持するため、突然変異率は通常 0.1~5%程度の低い値に設定される。

以上の交叉、突然変異、選択によって、個体集団を更新することを世代交代と呼び、この世代交代を繰り返すことによって最適化問題の解を求める。

GA では、染色体の情報を遺伝子型 (Genotype)、解の外部表現を表現型 (Phenotype) と呼ぶ。通常、遺伝子型は整数の配列で表現される。例えば、基本的な GA のコーディングの例では、遺伝子型は 2 ビットの 0 または 1 の整数による配列で表現される。この際に遺伝子型の長さを 4 とし、表現型は 4 ビットの整数で表現されるとすると、表現型は 0~16 の整数で表される。また、遺伝子型の染色体の情報に実数を用いた GA (Real-coded Genetic Algorithm; RCGA) に関する研究も行われている [70]。遺伝子型において実数値を扱う場合、その交叉処理に実数値用の手法が必要である。代表的な手法としては、BLX- α [71]、単峰性正規分布交叉 (Unimodal Normal Distribution Crossover; UNDX) [72]、シンプレクス交叉 (Simplex Crossover; SPX) [73]が提案されている。BLX- α は 2 個体の親によって定まる各変数の区間を α 倍に拡張してできる超直方体の子個体を一様乱数によって生成する交叉である。UNDX は 3 つの親個体によって決まる正規乱数を用いて子個体を生成する。SPX は n を次元数としたときに $n+1$ の親個体を頂点とする n 次元単体 (Simplex) に相似な単体内に一様分布に従って子個体を生成する交叉である。

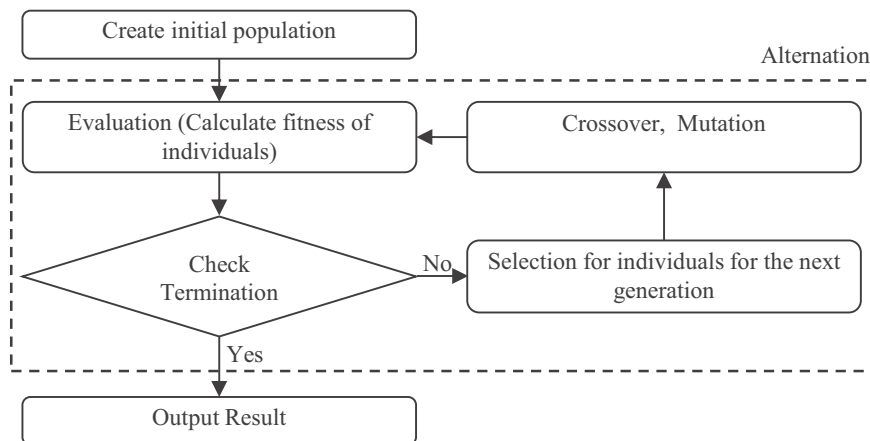


図 3.2 基本的な GA の処理の流れ

Fig. 3.2 Basic procedure of GA

3.2.3 遺伝的プログラミング (Genetic Programming; GP)

GP は 木構造の表現型を持つ GA を拡張した手法である。最適化における処理は GA と同様であり、選択、交叉、突然変異による世代交代を通して与えられた問題に適した木構造のモデルを自動生成する。木構造によって表現可能な問題として関数、手続き、プログラムおよび回路などがあることから、基本的な GA と比較して、処理アルゴリズムの自動構築において有効である。具体的には、最適な処理方法が未知の問題において、処理アルゴリズムが不明な場合、また処理アルゴリズムの自由度を高めたい場合に GP は有効である。GP では木構造を表現する方法として S 式²⁵が用いられ、木の節 (node) を演算子、木の葉 (leaf) を演算子の引数として数式やプログラムを構築する。図 3.3 は GP における解の一例である。図 3.3 の上に示した構造は数式を表現した例であり、図 3.3 の下に示した構造は論理演算を含めたプログラムを表現の例である。GP における進化では、GA と遺伝子型の構造が異なるため交叉の方法が異なる。GA に用いられるような交叉方法を適用すると、木構造における処理の手続きが破壊される可能性があるため、一般的には各個体が木であることを考慮して、各親個体の木の一部 (部分木) を交換する部分木交換交叉を行われる。また、木構造の最適化においては、GP によって進化を繰り返すうちにその構造が巨大になるブloat (bloat) が発生するという潜在的な問題がある。ブloat が発生すると木が冗長となり、解の探索が阻害される。ブloat を防ぐ単純な方法として、木の高さやノードの数を制限する方法が考えられるが、制約条件は個体の多様性を低下させることに繋がり、探索効率の低下が懸念される。ブloat を抑制する方法については研究が行われており、木の大きさに応じたペナルティ与える項を評価関数に加える方法や、遺伝操作を改良する

²⁵ プログラミング言語の Lisp で用いられ、式を (演算子 引数1 引数2) のように表現する方法。例えば、 $a * x$ は $(* a x)$ と表現される。

方法が提案され、ブロートの抑制において一定の効果を示している。しかし、ブロートに関する問題を完全に解消する最適な方法はないため、GP を最適化問題へ適用する際にはブロートに対応する方法の検討が必要である。

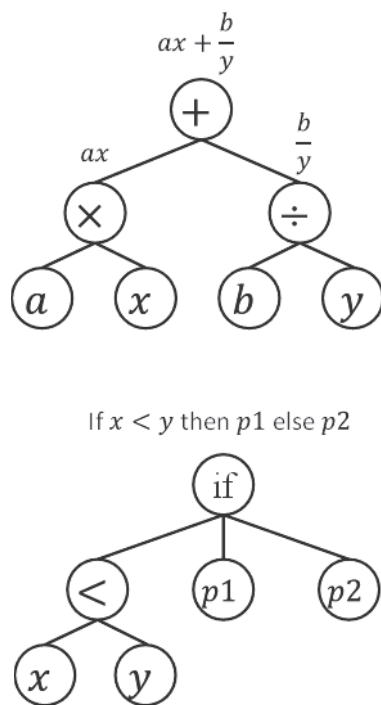


図 3.3 GP の表現型の例
Fig. 3.3 Example of Phenotype of GP

3.3 ニューラルネットワーク (Neural Network; NN)

神経回路網 (Neural Network; NN) は生体システムにおける脳の仕組みを模倣し、情報処理に適用することを目的として提案された手法である [74]. NN は人間の神経細胞にあたる要素を多数の相互接続された多入力 1 出力の演算ユニット (ニューロン) で構成される, ユニット間には実数値の結合荷重が存在し, あるユニットには, 接続された他のユニットから結合荷重に対応した重み付きの出力が入力される. ユニットにおける出力は接続された他のユニットからの入力の総和を入出力関数 f に代入した値である. NN では学習を通してユニット間の結合荷重を調節する. 具体的なシステムを構築する. 代表的な学習方法としては Rumelhart らによる誤差逆伝搬法 (Back Propagation method; BP method) [75] [76] が挙げられる. また, NN における回路網の構造は大きく分けて Rosenblatt よって提案された階層型 (a) と Hopfield ネットワーク [77] [78] を代表とした相互結合型 (b) の 2 種類に分類される. 図 3.4 に階層型と相互結合型の構造を示す. 階層型はユニットを複数の層状に配置し, 入力側から出力側の向きへ結合させた構造であり, 信号の伝達方向は決まっている. 主にパターン分類・識別などのモデルに用いられる. 相互結合型は各ユニットを相互に結合させたもので, 入出力の伝達方向は決まっておらず, 自己想起やエネルギー最小化のモデルに用いられる. NN は回路網の構造や結合荷重のパラメータの値を変更することで多様な処理に適用可能な汎用性の高いモデルであり, 様々な実用問題に適応されている. 本研究では第 6 章にて階層型 NN を用いた非線形回帰を用いた予測モデルの構築を行う. 階層型 NN における非線形回帰では, 入力と正解出力の組からなる学習用データ $\{x_p, u_p | p = 1, \dots, P\}$ に基づいて, 入力 $x \in R^l$ から望みの出力 u を推定する次の (3.1) 式のような非線形の変換をネットワーク上のユニット間の結合荷重を調整することによって構成する.

$$\hat{u} = \Phi(x) \quad (3.1)$$

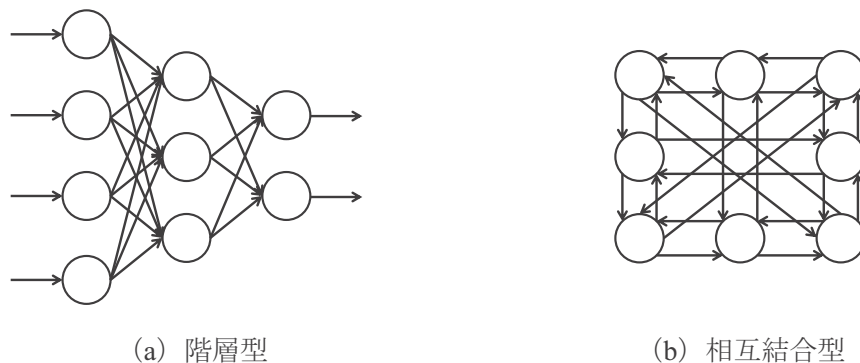


図 3.4 NN の構造

Fig. 3.4 Example of structure of NN

3.4 サポートベクターマシン (Support Vector Machine; SVM)

サポートベクターマシン (Support Vector Machine: SVM) [79] [80] は統計的学習理論 (Statistical learning theory) に基づく学習システムによって分類器を作成する手法である。単純パーセプトロン²⁶を用いて、2 クラスのパターン識別器を構成する手法である。サポートベクターマシンの基本的な 2 値分類問題における線形モデルは

$$y(x) = w^T \phi(x) + b \quad (3.2)$$

で表される。一般に学習によって構成された分類器には学習データに含まれていない未知データへの汎化性能が求められるが、SVM では「マージン」という概念を用いることで、未知のデータに対する汎化性能を向上させている。マージンとは図 3.5 に示すような分類境界と訓練データとの最短距離を示す。SVM ではこのマージンを最大化するという基準を置くことで分類境界を一意に定めている。さらに「カーネルトリック」と呼ばれるテクニックが提案されたことによって非線形な識別器の構成が可能となった。カーネルトリックとは学習アルゴリズムにおいて、入力ベクトル x が、スカラー積の中にのみ現れるとすれば、スカラー積をカーネル関数に置き換えることができるという考え方である。実用的なカーネル関数としては多項式カーネル、ガウスカーネル、シグモイドカーネルなどが挙げられる。SVM は、多くの識別手法の中でも識別性能の良い学習モデルの一つであり、画像分類、テキスト分類、手書き文字認識など、様々な現実的問題へ適用されている。

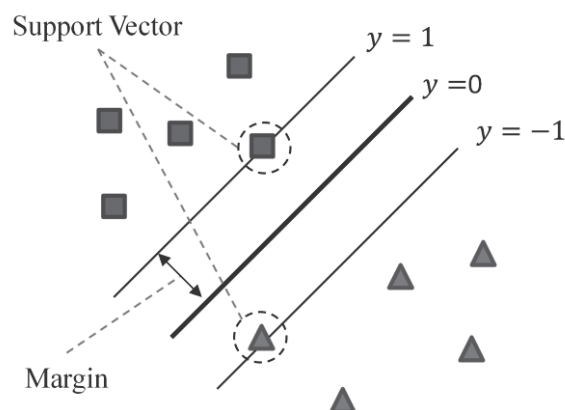


図 3.5 SVM におけるマージン

Fig. 3.5 Margin in SVM

²⁶ 神経回路網のモデルにおいて最も単純な線形しきい素子。

3.5 機械学習を用いた紙の製造に関する従来研究

紙の製造において、複雑な問題の予測やシミュレーションを行うための手段として機械学習を用いた研究が行われ、予測精度の向上や品質管理における良好な成果が報告されている。以下に機械学習を用いた紙の特性および製紙工程の予測に関わる従来研究をいくつか紹介する。

① GA による変数選択と多変量統計法による結合強さの予測

製造工程における製造条件の変動に関する複数の変数から製品における内部結合強さを予測するためのキャリブレーションモデルを、GA を用いて構築する研究 [81]が行われ、GA によって変数選択を行ったキャリブレーションが、最小二乗法およびニューラルネットワークと比較して高い予測性能を示し、実際の製造データに適用した場合にも良好な結果が得られていたことが報告されている。

② NN を用いた紙特性、木材特性、繊維特性の関係の予測モデルの構築

O'Neill らによって NN と重回帰分析を用いて、紙の特性と木材特性およびパルプにおける繊維の特性に関する予測モデルを構築する研究 [82]が行われ、NN を用いて構築されたモデルが重回帰分析を用いて構築したモデルより良好な予測結果を示したことが報告されている。

③ 多項式とニューラルファジーモデルによる紙の特性予測

代替の非在来型植物のパルプの製造において、多項式とニューラルファジーモデルを用いた紙の特性予測に関する研究 [83]が行われている。実験においては、複数の条件のパルプを作成し、異なる条件化で作製した紙の特性を予測するモデルを構築し、予測結果においてニューラルファジーモデルは依存変数について 8%から 10%の誤差で再現可能であったという良好な結果を得られたことが報告されている。

④ NN を用いた抄紙機上における断紙の要因解析

抄紙機上での断紙を改善するために NN を用いた研究 [84]が行われている。オンラインのウェットエンドセンサから得られる、ゼータ電位、歩留り、電導度および pH を測定し、断紙に関係する主要変数を 3 階層の NN と BP 法による学習を用いることによって抽出し、変数を基に対策を取ったところ、断紙回数が減少し、品質が向上するという良好な結果が得られたことが報告されている。

⑤ NN と GA を用いた紙の白色度の予測

製造工程において存在する多数の相互関係にある変数に影響される白色度を予測する手法として NN と GA を用いた研究 [85]が行われている。報告では、NN による予測モデルは 5%以内の誤差で白色度を予測可能で、さらに NN モデルに GA を組み合わせることによって変数の最適化が図られ、製造上のコストの低減に高い効果があったことが示されている。

⑥ SVM を用いた生産工程における変動の認識

自己相関性のある生産工程の管理においてトレンドの異常の誤検出を低減させるために SVM を用いた検証 [86]が行われている。報告では、工程中の異常でない状態を異常と誤認識する確率と、工程中に存在する異常を認識できない確率の最小化において SVM が高い効果を示したとされている。また、相関性のない工程の監視においても SVM は有効であったと報告されている。

以上のように紙の特性の予測や製造工程の管理における精度向上において機械学習を用いた研究は多数行われ、その有効性は示されている。しかしながら、本研究の目的であり、機械学習における課題の一つである、予測や判定結果の根拠や導出過程を説明する方法については別途検討が必要である。

第4章 特徴量選択型 SVM による紙の分類器の自動構築

4.1 はじめに

材料や製造条件等が異なることによって紙が持つ固有の特性は変化する，製造工程においてはこの特性を正確に把握し，分類することが必要とされている．紙の分類手法についてはこれまでに多くの手法が提案され，その手法の一つとして紙の画像に表れる特徴を用いることによって非破壊で紙を分類する方法に関する研究 [87] [88] [89] が行われており，その有効性が示されている．これらの手法は，製造工程に由来する紙の特性の変動を撮像した画像に対して周波数解析を行いうことによって解析し，その解析結果を基に分類を行う手法であり，製造工程の異なる紙を分類する方法として有効であると報告されている．一方，画像解析においては，様々な特徴量算出方法があり，それらを通じて撮像対象における様々な情報を取得することが可能である．画像から得られる特徴量において，単一の特徴量だけでなく複数特徴量を用い，それらを組み合わせた分類器を作成することによって，紙の分類能力を向上させる可能性が考えられる．近年，画像を分類する際，画像解析によって様々な特徴量を算出し，NN, SVM 等，機械学習によって分類器を構成する手法 [90] [91]の応用例や有効性が報告されている．これらの手法は，複数の特徴量からなる多次元の特徴量空間において境界となる非線形の識別面を作成することから，様々な現実的な分類問題に適用することが可能であり，製造工程の条件によって様々な特徴が変化する紙の分類においても有効であると考えられる．しかし，大量の特徴量をそのまま学習した場合，計算コストの増大や，有効でない特徴量の影響による分類精度の低下などの問題点がある．そのため，有効な特徴量だけを選択し，次元数を削減することが課題である．この課題における従来の方法として，主成分分析などの多変量解析の手法が用いられているが，特徴量間に強い相関関係がある場合はあまり有効ではない．また，特徴量数が多い場合，複数の特徴量の中から，分類結果の精度を考慮しながら有効な特徴量だけを人間の手によって発見することは容易ではない．さらに，分類する対象によって有効な特徴量が異なることが予想されるが，分類器ごとに人間の手で特徴量の選択を行うことは，分類精度を不安定にする要因になると考えられる．そこで，本章では GA による特徴量の組み合わせの最適化と SVM による学習を組み合わせた分類器作成手法として提案されている特徴量選択型 SVM (Feature selection type SVM; FS-SVM) [92]を用い，画像解析によって算出可能な 239 種類の特徴量の組合せ最適化を行うことによって，高精度かつ分類に必要な特徴量を明確に示すことが可能な紙の分類器の作成方法を検討する．

4.2 特徴量選択型 SVM (Feature selection type SVM; FS-SVM)

画像を分類する際、画像から大局的または局所的な複数の特徴量を算出し、それらの高次元なデータを基に学習を行い、分類器を構成する手法が多く用いられている。しかし、画像分類で扱うような高次元ベクトルをそのまま学習した場合、計算コストの増大やノイズの影響が懸念される。そこで、GA を用いて最適な特徴量の組み合わせを選択し、選択した特徴量のみを用いて SVM を学習させ、分類器を自動的に構成する方法である特徴量選択型 SVM が提案されている。本手法においては、画像を用いた欠陥製品の分類問題において、全ての特徴量を入力した従来の SVM と比較した結果、未知画像に対する正答率が向上したことが余部らによって報告されている。特徴量選択型 SVM における分類器の構成手順を図 4.1 に示す。始めに、学習画像および評価画像から特徴量を算出する。学習画像とは、SVM の学習に用いる画像であり、評価画像とは、GA における適応度を計算する際に用いる画像である。次に、初期個体群を生成する。ここで、各個体の染色体は図 4.2 のように特徴量の使用の有無を 0, 1 のビット列で表現するものとする。次に、学習画像において選択された特徴量を用いて構成された分類器を評価画像に適用し、その結果を基に GA における適応度を算出する。以上の手順の繰り返しによって、特徴量の組み合わせを最適化した分類器を構成する。

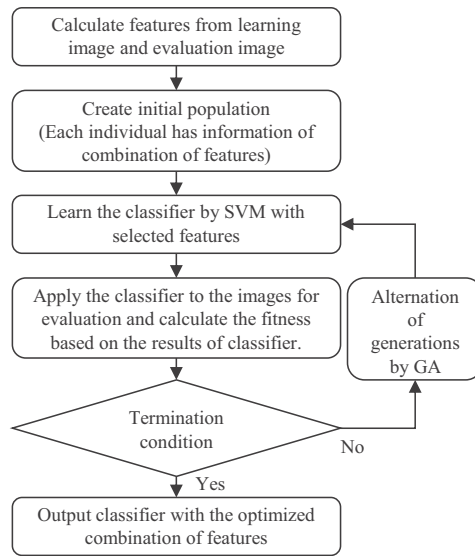


図 4.1 特徴量選択型 SVM における最適化手順

Fig. 4.1 Optimization procedure of FS-SVM

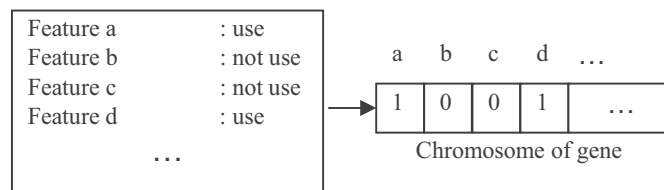


図 4.2 特徴量選択型 SVM における遺伝子型

Fig. 4.2 Genotype of FS-SVM

4.3 実 験

市販の印刷用紙および作製条件の異なる手すきシートをサンプルシートとし、特徴量選択型 SVM よって各テストシートの分類器を構成する実験を行った。

4.3.1 テストシート

① メーカーの異なる PPC 用紙²⁷

テストシートとしてメーカーの異なる 5 種類の PPC 用紙を用いた。テストシート名はテストシート A, テストシート B, テストシート C, テストシート D, テストシート E とする。坪量が約 80g/m² の紙を用いた。

② 製造条件の異なる手すきシート

一般の PPC 用紙とは異なるテストシートとして、製造条件の異なる手すきシートを作製した。変更した条件は、紙料の種類、叩解回数、てん料の添加率、定着助剤の添加率、ワイヤーの種類である。紙料の種類は、広葉樹晒パルプ (LBKP)、針葉樹晒クラフトパルプ (NBKP) およびアバカパルプ (ABACA) を用い、叩解には PFI ミルを使用した。てん料は白土、定着助剤は硫酸アルミニウムとし、ワイヤーの種類は網目の異なるブロンズ製ワイヤー 3 種類 (ブロンズ 1 目開き: 縦 106 μ m, 横 106 μ m, ブロンズ 2 目開き: 縦 294 μ m, 横 177 μ m, ブロンズ 3 目開き: 縦 160 μ m, 横 198 μ m) とプラスチック製ワイヤー 1 種類 (日本フィルコン社製 SS-400) を用いた。各種サンプルの坪量は絶乾で 60g/m² とし、作製に当たっては JIS P8222 に則った。表 4.1 に各手すきシートのテストシート名および各条件等を示す。各条件における手すきシートはそれぞれ 3~5 枚作製し、シワやチリなどが少ないものを各水準 2 枚または 3 枚抜粋した。

²⁷ 普通紙複写機用の印刷用紙の事 (Plain Paper Copier; PPC)。

表 4.1 手すきシートサンプルの作製条件

Table 4.1 Settings for making hand sheet

Name of sample	Number of beating	Freeness (mL CSF)	Ratio of fixing agent (%)	Ratio of filler (%)	Type of wire	Basis weight (g/m ²)
LBKP-1	0	670	0	0	Bronze 1	66.6
LBKP-2	4000	430	0	0	Bronze 1	66.9
LBKP-3	4000	430	1	0	Bronze 1	67.5
LBKP-4	4000	430	2	0	Bronze 1	67.8
LBKP-5	4000	430	1	2	Bronze 1	67
LBKP-6	4000	430	1	5	Bronze 1	68.7
LBKP-7	8000	360	0	0	Bronze 1	65.8
LBKP-8	8000	360	0	0	Bronze 2	65.5
LBKP-9	8000	360	0	0	Bronze 3	65.8
LBKP-10	8000	360	0	0	Plastic	65.5
NBKP	0	660	0	0	Bronze 1	67.5
ABACA	0	650	0	0	Bronze 1	66.7

4.3.2 テストシートの撮像方法

実験に用いるテストシートの画像には、各テストシートの透過画像を近赤外波長領域で撮像したものを採用することとした。赤外光を用いた理由としては、可視光に比べて透過率が高いことから汚れ等のノイズ成分の低減が可能になるためである。撮像条件は、輝度分解能 256 階調、空間分解能 400dpi (1 画素=63.5 μ m 角) とし、測定時には光源側がワイヤー接触面となるように統一した。測定装置にはアイメジャー社製の赤外透過型撮像装置 (IR-6000) を用いた (図 4.3)。図 4.4 に撮像した各テストシートの画像の一例として、作製した手すきシートのうち、条件が大きく異なる LBKP-1, LBKP-4, LBKP-6, LBKP-7, LBKP-8, LBKP-9, LBKP-10, NBKP, ABACA の画像を抜粋して示す。なお、図 4.4 に記載の画像は紙面上での視認性を考慮し、画像処理によって階調を補正した画像である。

また本実験では、測定画像から縦横 200 画素 (約 12.5mm 角) の領域を 1 枚の紙片とする紙片画像を手すきシートごとに複数枚作成し、それらの各画像に対して学習および評価を行った。この際、各画像群において紙片の切出し位置が偏らないようにテストシートの全体画像からランダムに切り出した。



図 4.3 赤外透過画像撮像装置

Fig. 4.3 Transparency NIR (Near Infra-Red) scanner

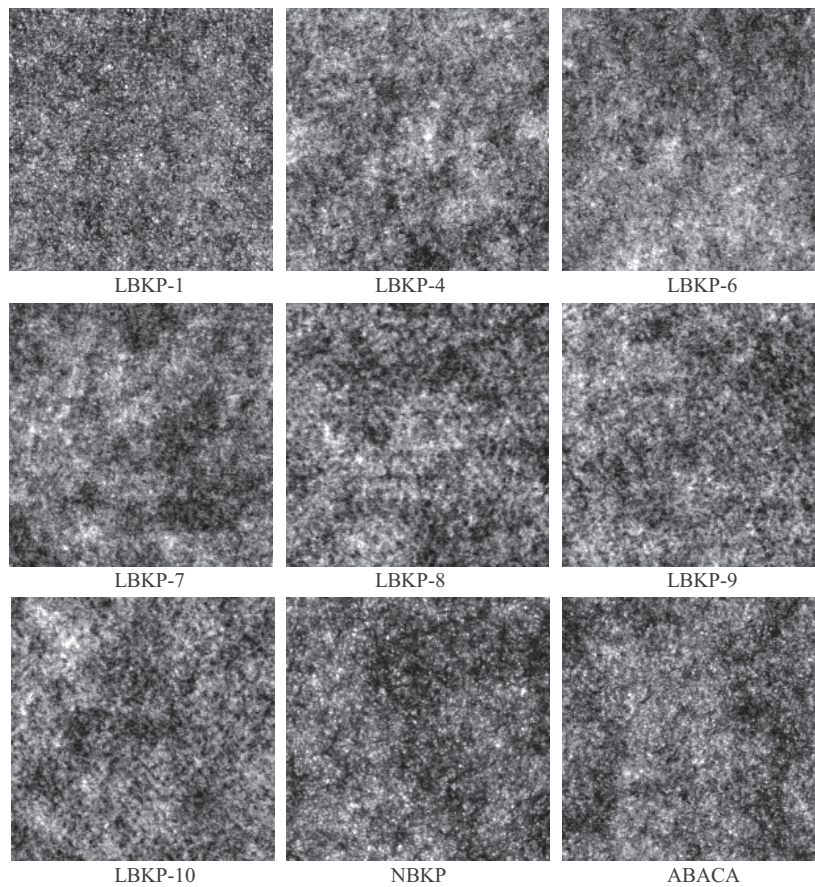


図 4.4 各サンプルの赤外透過画像撮像

Fig. 4.4 Infra-red transmission image of each sample

4.3.4 画像特徴量の算出

分類に使用する特徴量として、本研究では画像から算出可能な以下に記載する特徴量を用いた。表 4.2 に特徴量の一覧を示す。各特徴量は紙片画像ごとに算出した。

① 画像の階調値における統計量

対象画像の階調値から得られる特徴量。記号を a-1 から a-17 とする。

② 2 値画像から得られる特徴量

2 値化した対象画像から得られる特徴量。2 値化には判別分析法 [93]を用いた。記号を b-1 から b-16 とする。

③ 分割した画像から得られる特徴量

対象画像を 2 つに分割して得られる特徴量。画像の対称性を考慮するため、1 枚の画像を上下(UD)、もしくは左右(LR)2 つに分割して、2 組の画素対の統計解析を行い、その値を特徴量とした。記号を c-1 から c-3 とする。

④ Gray Level Co-occurrence Matrix (GLCM) から得られる特徴量

テクスチャ解析で用いられる手法として、Gray Level Co-occurrence Matrix(GLCM) [94]を用いた画像解析方法が提案されている。上記手法は、画像における画素の濃度の一様性、方向性、コントラストの変化などの特徴を求める手法である。GLCM は、画像において、ある画素 i に注目した際に、任意に指定する 2 つのパラメータである距離 d と角度 θ に対応する画素 j との濃度の関係を行列 $p(i, j | d, \theta)$ に表した確率行列である。GLCM を作成するに当たってパラメータである距離 d および角度 θ の指定が必要であるが、本研究における実験では、距離 d を 1~8、角度 θ を 0° 、 45° 、 90° 、 135° および 4 方向全ての 5 種類とした。上記パラメータの組み合わせで作成される GLCM は全部で 40 通りとなる。各 GLCM から算出する特徴量はそれぞれ異なる特徴量とした。記号を d-1 から d-5 とし、各パラメータ d および θ は、記号の末尾に記載する。 θ については 0° 、 45° 、 90° 、 135° および 4 方向全てを、それぞれ 0~4 とする。(例：パラメータ $d=1, \theta=90^\circ$ における Energy の場合、d-1_1_2 と記載する)

表 4.2 画像特徴量一覧
Table. 4.2 List of image feature

Features calculated from the pixel values		Features obtainable from the binary image		Features obtainable from the segmented image		Features computable from the GLCM	
Sign	Name of feature	Sign	Name of feature	Sign	Name of feature	Sign	Name of feature
a-1	Mean	b-1	Number of area	c-1	Covariance	d-1	Energy
a-2	Max	b-2	Total area	c-2	Correlation coefficient	d-2	Entropy
a-3	Min	b-3	Mean area	c-3	Kolmogorov-Smirnov static	d-3	Contrast
a-4	Difference between Min and Max	b-4	Max area			d-4	Homogeneity
a-5	Standard deviation	b-5	Mean contour length			d-5	Correlation
a-6	Mode	b-6	Max contour length				
a-7	In 1 Sigma	b-7	Feret diameter mean ratio of box area				
a-8	Out 3 Sigma	b-8	Feret diameter max ratio of box area				
a-9	Median	b-9	Difference ratio of horizontal line				
a-10	First Quartile	b-10	Difference ratio of vertical line				
a-11	Third Quartile	b-11	Mean degree of circularity				
a-12	Kurtosis	b-12	Max degree of circularity				
a-13	Skewness	b-13	Min degree of circularity				
a-14	Threshold of discriminant Analysis	b-14	Mean angle (Inertial equivalent ellipse)				
a-15	Between-class Variance	b-15	Max angle (Inertial equivalent ellipse)				
a-16	Lower within-class Variance	b-16	Min angle (Inertial equivalent ellipse)				
a-17	Higher within-class Variance						

4.3.5 特徴量選択型 SVM に関するパラメータ

特徴量選択型 SVM におけるパラメータとして、GA に関する条件、SVM の学習に関する条件の指定がそれぞれ必要である。各パラメータの設定内容を表 4.3 に示す。これらの値は予備実験によって適切な値になるよう調整して決定した。評価関数 F を以下の (4.1) 式に示す。

$$F = N_c + M_{ave} \quad (4.1)$$

N_c は正しく分類された数であり、 M_{ave} は SVM におけるマージンの平均値を示す。

表 4.3 特徴量選択型 SVM の学習パラメータ

Table. 4.3 Parameter of Learning of FS-SVM

(Conditions of GA)	
Number of Individuals	100
Generation alternation model	MGG
Method of Selection	Tournament selection
Method of Crossover	Uniform Crossover
Crossover Rate	1
Mutation Rate	0.05
Fitness function	Correct answer rate + SVM's average margin
Termination condition	20000 generations
(Conditions of SVM)	
Kernel Function	Gaussian kernel
Constant of Kernel Function	0.6
Learning coefficient	0.2
Number of learning	3000
Number of features	239

4.3.6 分類対象の設定

作成する分類器の対象を以下の Test1 から Test6 のように設定した (表 4.4). 各 Test は 2 分類問題として設定した. 表 4.5 に各 Test における学習画像, 評価画像, 未知画像の枚数を示す. また, 表 4.6 に各分類器における正答の定義を示す.

① Test1 : メーカーの異なる紙

メーカーの異なる紙であるテストシート A からテストシート E を対象とし, 前記 5 種類のテストシートが含まれる集団から特定のテストシートだけを抽出する分類器の学習を行う. テストシート A からテストシート E を抽出する分類器をそれぞれ 1-1, 1-2, 1-3, 1-4, 1-5 とする.

② Test2 : 原料の異なる手すきシートの分類

原料の異なる手すきシート LBKP-1, NBKP, ABACA を対象とし, 前記 3 種類のテストシートが含まれる集団から特定のテストシートだけを抽出する分類器の学習を行う. LBKP-1, NBKP, ABACA を抽出する分類器をそれぞれ 2-1, 2-2, 2-3 とする.

③ Test3 : 紙料叩解回数の異なる手すきシートの分類

紙料の叩解回数の異なる手すきシート LBKP-1, LBKP-2, LBKP-7 を対象とし, 各テストシート間の分類を行う分類器の学習を行う. LBKP-1 と LBKP-2 を分類する分類器 3-1, LBKP-2 と LBKP-7 を分類する分類器 3-2, LBKP-1 と LBKP-7 を分類する分類器 3-3 を作成する.

- ④ Test4：定着助剤の添加率の異なる手すきシートの分類
定着助剤の添加率の異なる手すきシート LBKP-2, LBKP-3, LBKP-4 を対象とし, 各テストシート間の分類を行う分類器の学習を行う. LBKP2 と LBKP-3 を分類する分類器 4-1, LBKP-3 と LBKP-4 を分類する分類器 4-2, LBKP-2 と LBKP-4 を分類する分類器 4-3 を作成する.
- ⑤ Test5：てん料の添加率の異なる手すきシートの分類
てん料の添加率の異なる手すきシート LBKP-2, LBKP-5, LBKP-6 を対象とし, 各テストシート間の分類を行う分類器の学習を行う. LBKP2 と LBKP-5 を分類する分類器 5-1, LBKP-5 と LBKP-6 を分類する分類器 5-2, LBKP-2 と LBKP-6 を分類する分類器 5-3 を作成する.
- ⑥ Test6：ワイヤーの異なる手すきシートの分類
ワイヤーの異なる手すきシート LBKP-7, LBKP-8, LBKP-9, LBKP-10 を対象とし, 前記4種類のテストシートが含まれる集団から特定のテストシートだけを抽出する分類器の学習を行う. LBKP-7, LBKP-8, LBKP-9, LBKP-10 を抽出する分類器をそれぞれ 6-1, 6-2, 6-3, 6-4 とする.

表 4.4 各 Test における分類対象

Table. 4.4 Target of classification in each test

		Target	Not Target
Test 1	1-1	A	Not A
	1-2	B	Not B
	1-3	C	Not C
	1-4	D	Not D
	1-5	E	Not E
Test 2	2-1	LBKP-1	Not LBKP-1
	2-2	NBKP	Not NBKP
	2-3	ABACA	Not ABACA
Test 3	3-1	LBKP-1	LBKP-2
	3-2	LBKP-2	LBKP-7
	3-3	LBKP-1	LBKP-7
Test 4	4-1	LBKP-2	LBKP-3
	4-2	LBKP-3	LBKP-4
	4-3	LBKP-2	LBKP-4
Test 5	5-1	LBKP-2	LBKP-5
	5-2	LBKP-5	LBKP-6
	5-3	LBKP-2	LBKP-6
Test 6	6-1	LBKP-7	Not LBKP-7
	6-2	LBKP-8	Not LBKP-8
	6-3	LBKP-9	Not LBKP-9
	6-4	LBKP-10	Not LBKP-10

表 4.5 各 Test で用いた画像枚数

Table. 4.5 Number of images used in each test

	Learning	Evaluation	Unknown
Test 1	75	75	30
Test 2	153	144	144
Test 3	102	96	96
Test 4	102	96	96
Test 5	102	96	96
Test 6	136	128	128

表 4.6 各分類器における正答の定義

Table. 4.6 Definition of correct answers

		Output of classifier	
		Target	Not Target
Correct	Target	$N1$	$N2$
	Not Target	$N3$	$N4$

4.3.7 比較手法

紙の画像を用いた紙の分類手法として、フーリエ変換と相互相関法による紙の異同識別法 [87] [88]が提案されている（以下、PS 法とする.）。本手法は、紙の画像に対して 2 次元フーリエ変換を施し、得られたパワースペクトル（PS）を用いる（図 4.5）。判定処理においては、まず基準となる紙のデータを収集し、各 PS 画像を作成する。その後、各 PS 画像の平均画像を作成し参照 PS 画像とする。そして判定時においては基準となる参照 PS 画像と判定対象の紙から作成した PS 画像の類似度を相互相関法によって算出し、その値を基に紙の異同識別を行う方法であり、紙の画像が持つ周期的な特性に注目した分類方法である。本手法を比較手法として、実験に用いたテストシートの画像に対してこの手法を適用した。適用に際して、参照 PS 画像は表 4.5 における各学習画像群における PS 画像の平均画像とした。また、PS 画像において、中心から 10 画素の PS 値は 0 とした。フーリエ変換処理は各画像における画素数が縦横 200 画素であることから離散フーリエ変換（Discrete Fourier Transform; DFT）を用いて行った。

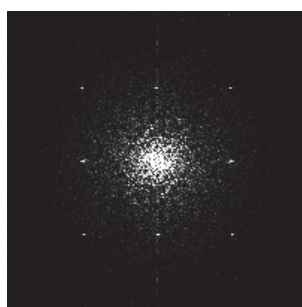


図 4.5 ワイヤーの特徴が現れている紙の PS 画像の一例

Fig. 4.5 Example of PS image of paper which shows the characteristic of wire

4.4 結果

4.4.1 各分類器の分類正答率

各 Test で構成された分類器における学習過程における最良個体の適応度の推移を図 4.6, 評価画像および未知画像に対する分類結果を表 4.7 に示す. 表における正答率は (4.1) 式で表され, 評価画像または未知画像の総枚数に対して正しく分類した画像の枚数を割合 (Correct Answer Rate; CAR) で示したものである. $N1, N2, N3, N4$ の各記号は表 4.6 に対応する.

$$CAR = \frac{N1+N4}{N1+N2+N3+N4} \quad (4.2)$$

各 Test における分類器の作成回数は 10 回とし, 作成された分類器のうち, 未知画像に対する分類正答率が最良のものを結果として用いた. 評価画像に対しては, Test4 の 4-3, Test5, Test6 の 6-3, 6-4 を除いて 100% の正答率となった. 未知画像に対しては, Test1, Test2, Test3 においては 100%, それ以外にも Test5 の 5-2 を除いてほぼ 90% 以上の正答率であった. 図 4.6 の推移は 10 回の試行において最良結果を得られた分類器が作成された際の推移である. また, 未知画像に対する分類正答率が 100% の分類器においては最大適応度の推移の様子が類似していることから代表して 1-2 のみを示す. その他の分類器においては分類正答率が 100% に満たなかった分類器である 4-3, 5-1, 5-2, 5-3, 6-3, 6-4 を抜粋して示す.

4.4.2 選択された特徴量

各分類器において選択された特徴量を表 4.8 に示す. 選択された特徴量の種類に注目すると, 階調値を用いた特徴量(a 群)と GLCM から得られる特徴量(d 群)が多く選択されていることが確認でき, それぞれ, 他の特徴量群との組み合わせによって構成されている. 1-1 と 3-1 だけは単一の特徴量群で構成されており, 1-1 においては a 群だけ, 3-1 においては d 群だけでそれぞれ構成されていることを確認することができる. 分割した画像から得られる特徴量 (c 群) については 2-2, 5-2 を除いて選択されなかった.

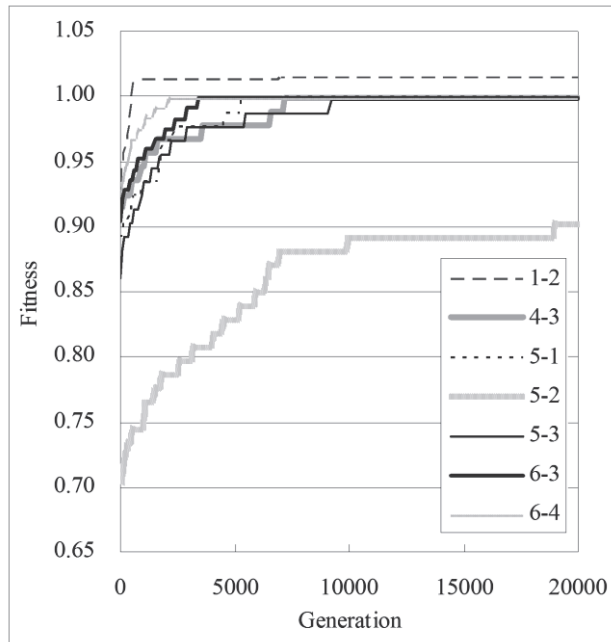


図 4.6 FS-SVM の学習における最良個体の推移

Fig. 4.6 Transition of maximum fitness in learning of FS-SVM

表 4.7 FS-SVM による各 Test における分類正答率

Table 4.7 Correct classification ratio using FS-SVM in each Test

Test 1	1-1	1-2	1-3	1-4	1-5
Evaluation (%)	100	100	100	100	100
Unknown (%)	100	100	100	100	100
Test 2	2-1	2-2	2-3		
Evaluation (%)	100	100	100		
Unknown (%)	100	100	100		
Test 3	3-1	3-2	3-3		
Evaluation (%)	100	100	100		
Unknown (%)	100	100	100		
Test 4	4-1	4-2	4-3		
Evaluation (%)	100	100	99		
Unknown (%)	100	100	96.9		
Test 5	5-1	5-2	5-3		
Evaluation (%)	99	89.6	99		
Unknown (%)	89.6	66.7	92.7		
Test 6	6-1	6-2	6-3	6-4	
Evaluation (%)	100	100	99	99	
Unknown (%)	100	100	93.8	91.7	

表 4.8 FS-SVM によって各分類器に選択された特徴量

Table 4.8 Selected features in each classifier with FS-SVM

		Selected features	Group				Total
			a	b	c	d	
Test 1	1-1	a-1,a-6,a-9,a-11,a-14,a-16	6	0	0	0	6
	1-2	a-1,a-10,a-11,a-14,d-3_1_1	4	0	0	1	5
	1-3	a-1,a-3,a-4,a-11,a-14,b-13,d-1_0_1,d-2_1_4,d-2_2_5,d-5_0_2,etc	5	1	0	10	16
	1-4	a-1,a-6,a-11,d-5_1_2,d-5_4_2,d-4_0_2,d-3_0_1,d-3_0_2	3	0	0	5	8
	1-5	b-8,d-1_4_3,d-2_2_5	0	1	0	2	3
Test 2	2-1	a-1,a-10,a-15,a-16,b-13,d-1_2_6,d-2_0_3,d-2_0_5,d-2_0_8,d-2_1_3,etc	4	1	0	24	29
	2-2	a-16,a-17,c-3_LR,b-13,b-15,d-1_1_2,d-2_1_5,d-2_4_1,d-2_4_4,d-5_1_7,etc	2	2	1	12	17
	2-3	a-11,a-14,a-16,b-16,d-1_0_8,d-1_1_3,d-2_0_1,d-5_1_8,d-4_1_3,d-4_1_5,etc	3	1	0	15	19
Test 3	3-1	d-2_2_6,d-5_1_7,d-5_3_1,d-5_3_2,d-4_0_2,d-3_1_3,d-3_3_1,d-3_4_3	0	0	0	8	8
	3-2	a-1,a-14,d-5_3_8,d-4_3_3,d-3_0_1,d-3_1_6,d-3_1_8	2	0	0	5	7
	3-3	a-1,a-2,a-9,a-10,d-5_0_1,d-5_3_1,d-4_2_1,d-3_3_1	4	0	0	4	8
Test 4	4-1	a-1,a-6,a-9,a-10,a-14,d-5_4_4	5	0	0	1	6
	4-2	a-1,a-6,a-9,a-10,a-14,d-1_3_4,d-2_1_2	5	0	0	2	7
	4-3	a-9,b-1,d-5_0_3,d-5_1_2,d-4_0_1,d-4_1_2,d-3_0_4,d-3_1_2,d-3_3_2	1	1	0	7	9
Test 5	5-1	a-5,a-15,b-1,b-11,d-5_0_4,d-5_0_5,d-5_1_8,d-5_3_7,d-4_0_5,d-4_3_1,etc	2	2	0	16	20
	5-2	a-12,a-16,a-17,c-1_UD,c-2_UD,b-7,b-12,b-10,d-1_0_4,d-1_0_8,etc	3	3	2	32	40
	5-3	a-13,b-1,b-6,b-7,b-12,b-13,d-5_1_8,d-5_2_5,d-5_3_5,d-4_0_1,etc	1	5	0	19	25
Test 6	6-1	a-13,d-2_2_5,d-4_2_4,d-4_3_8,d-3_1_5,d-3_1_6,d-3_2_8,d-3_3_6,d-3_4_8	1	0	0	8	9
	6-2	a-1,a-2,a-12,a-13,d-2_3_8,d-4_0_2,d-4_2_2,d-4_3_1,d-4_4_1,d-3_0_2,etc	4	0	0	11	15
	6-3	a-1,b-7,d-1_0_3,d-1_2_2,d-1_4_3,d-5_0_2,d-5_0_3,d-5_0_4,d-5_0_5,d-5_0_7,etc	1	1	0	36	38
	6-4	a-2,a-12,a-13,b-13,d-2_0_8,d-5_0_2,d-5_0_4,d-5_0_5,d-5_0_7,d-5_1_1,etc	3	1	0	22	26

4.5 考 察

4.5.1 従来手法との比較

従来手法である PS 法による分類結果を表 4.9 に示す。表 8 における正答率は未知画像の総枚数に対して、正しい分類結果を得られた枚数の割合を示したものである。表 4.7 と表 4.9 から本研究で用いた手法はメーカーの異なる紙および作製条件の異なる手すきシートの分類において高い性能を有していることを確認することができる。本研究における提案手法の結果を比較手法の結果と比較すると、メーカーの異なる紙およびワイヤーの異なる手すきシートの分類においては同等の分類性能、その他のテストシートの分類では Test5 を除いて、提案手法において安定した高い性能を確認することができる。この結果から提案手法の有効性を確認することができた。

表 4.9 フーリエ変換と相互相関法による紙の異同識別法による分類正答率

Table 4.8 Correct classification ratio based on Fourier transform and cross-correlation method

Test 1	1-1	1-2	1-3	1-4	1-5
Correct (%)	100	100	100	100	100
Test 2	2-1	2-2	2-3		
Correct (%)	36.7	100	100		
Test 3	3-1	3-2	3-3		
Correct (%)	99	64.3	74.5		
Test 4	4-1	4-2	4-3		
Correct (%)	67.3	61.2	40.8		
Test 5	5-1	5-2	5-3		
Correct (%)	42.9	99	76.5		
Test 6	6-1	6-2	6-3	6-4	
Correct (%)	100	100	100	100	

4.5.2 てん料の添加率の異なる紙の分類

表 4.7 から、てん料の添加率の異なる手すきシートの分類器である 5-2 において、評価画像および未知画像に対する分類正答率が他の分類器に比べて低いことがわかる。また、図 4.6 から適応度の上昇も他の分類器と比較して遅いことがわかる。原因として今回作製した手すきシート LBKP-5 (てん料添加量 2%) と LBKP-6 (てん料添加量 5%) の間においては撮像画像に差が現れなかった可能性が考えられる。また、本実験で用いた特徴量の中に分類するために有効な特徴量が存在しない可能性も同時に考えられる。したがって、てん料の異なるテストシートの分類に関しては、テストシートの測定方法の変更、または本実験で使用しなかった特徴量の追加などを検討することが必要である。

4.5.3 分類に有効な特徴量

各分類器において選択された特徴量においては、階調値を用いた特徴量と GLCM を用いた特徴量が多く選択された。階調値のみまたは GLCM の特徴量のみによる分類器も構成され、それぞれ高い分類正答率を示したことから、各特徴量が紙の製造条件の違いによって現れる画像上の特徴量に反映され、紙の分類に有効に働いたと考えられる。また、それぞれの Test における分類器で使用されている特徴量は類似するものが多いが、その組み合わせはそれぞれ異なった。これは、各 Test における画像の特徴が、条件の違いによって複雑に変化したことが要因になっていると考えられる。また、複数の組み合わせの相互作用によって、異なる特徴量を使用しつつも、同様の効果をもった分類器が作成されたことも考えられる。このような判断を人の手で行うことは困難であることから、学習によって自動で有効な特徴量を選択し分類器を構成する方法が有効であったと考えられる。

4.5.4 特徴量の追加

特徴量選択型 SVM は、複数の特徴量の中から有効でないと考えられる特徴量を削減する機能を有することから、新たな特徴量を追加した際にも、有効でない特徴量が SVM における学習の際に使用されることを抑制することが期待できる。よって、本実験において使用していない画像特徴量、または画像以外の特徴量としてその他の測定方法から得られる特徴量も合わせて用いることが有効である。

4.6 まとめ

製造条件の異なる紙の分類において、画像から算出した複数の画像特徴量の最適化と SVM によって、自動的に紙の分類器を構成する実験を行った。その結果、メーカーの異なる紙およびワイヤーの異なる手すきシートの分類に加えて、その他の様々な条件が異なる手すきシートに対しても高精度な分類器を自動的に構成することができ、提案手法の汎化性能と有効性を確認することができた。本手法の各変動要因に対する識別能については、さらに詳細な水準を設定したテストシートを用いた実験による検証が必要と考えており、今後の課題として検討する予定である。また、実験に用いたテストシートのうち、てん料の添加量を変化させた手すきシートの分類においては、他のテストシートにおける分類結果と比較して分類正答率が低い結果となった。この結果を受け、その他の測定方法によって得られる特徴量を利用し、識別に有効な特徴量を見出すことが今後の課題である。

第5章 Evolutionary Decision Network を用いた紙の分類器の自動構築

5.1 はじめに

第4章では、特徴量選択型 SVM によって従来手法より分類精度が高く、また必要な特徴量が明確となった分類器を作成することができた。しかし、分類過程においては SVM を使用しているため、高い分類精度を持つ一方、その過程は複雑で人間が理解し易いとは言えない。そこで、必要な特徴量を示すだけでなく、分類に至る過程をより容易に理解可能な手法を検討する必要がある。人間が理解し易い過程を構築可能な分類モデルとしては決定木 [95] が挙げられる、決定木では単純な条件判断のみを繰り返すことによって分類モデルを構築する。しかし、決定木は複雑な問題においてはそのモデルが巨大なものとなる。また、モデルを構築する際にブロートが発生し、学習の効率が低下するとともに得られる分類器のモデルが不必要に大きくなることによって性能が低下するという問題がある。この課題において、進化的条件判断ネットワーク (Evolutionary Decision Network: EDEN) [96] が提案されている。本手法は単純な条件分岐ノードによってネットワーク状に構成し、入力したデータを分類するモデルを構築する。EDEN における各ノードは入力されたデータ群における単一の特徴量の値に注目した条件分岐のみを行う。モデルの構築においては、ノード間の接続、分岐における条件、各ノードが注目する特徴量などの組合せを進化計算によって最適化する。類似した処理を行う決定木と比較するとネットワーク構造を用いていることからシンプルなモデルを構築可能であるという報告がなされている。したがって、人に理解し易い分類過程をもつ分類器を構成することが期待できる。

本章では EDEN を用い、人間が理解し易い分類過程を持つモデルの構築に関する検討および実験を行った結果について述べる。

5.2 Evolutionary Decision Network (EDEN)

5.2.1 EDEN の概要

EDEN は条件分岐を行う条件判断ノードを組み合わせたネットワーク状の分類器を、進化計算法を用いた最適化によって構成する手法である。条件判断ノードにおける分岐は入力データにおける単一の特徴量の値にのみよって行われる。分岐の基準、分岐に用いる特徴量およびノード間の接続状態を進化計算によって最適化することによって目的の分類器を自動的に構築する。EDEN におけるネットワークは決定木の木構造を包含しており、また、決定木において問題となるブロートが発生しないことから、より人間に理解し易い分類過程を構築することが可能である。

5.2.2 EDEN の基本構造

EDEN の構造の例を図 5.1 に示す。提案されている EDEN の表現型は NN のようなフィードフォワード型のネットワークである。EDEN に用いられるノードは大きく分けて入力ノード、中間ノード、出力ノードの 3 種類に分類される。入力ノードは、入力されるデータ群の各経路に対応する。出力ノードは、分類器における分類クラスに対応する。中間ノードは、中間ノードはしきい値、分岐方法、分岐先、入力データにおける参照する特徴量のなどパラメータを保持しており、そのパラメータに従って条件分岐を行う。分類の主な流れは、入力ノードからデータが入力され、中間ノードにおいて再帰的な条件分岐が行われ、最終的に到達した出力ノードが分類されたクラスとなる。分岐においては、多様な方法に対応する。例えば、2 つ以上のしきい値を用いることによって、3 分岐以上やしきい値の範囲を表現することが可能である。

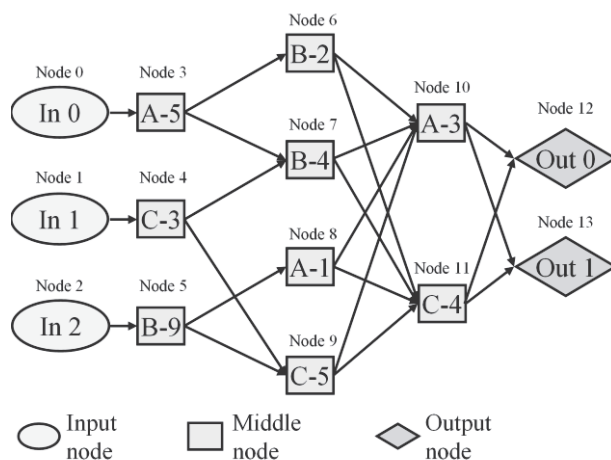


図 5.1 EDEN の構造の例

Fig. 5.1 Example of a structure in EDEN

5.2.3 EDEN における相互作用の表現

EDEN の特徴として多入力への対応が挙げられる。この特徴によってデータ群同士の相互作用の表現を実現可能としている。EDEN における多入力では、モデルにおける各ノードのしきい値を異なるデータ群の値を基に切り替えることで相互作用を表現する。図 5.2 を基に処理の流れの例を述べる。図 5.2 において、In 0 を注目している入力とし、In 1 および In 2 をその他のデータ群（参照入力）とする。まず、モデルのネットワーク内へ In 1 および In 2 を入力する。その際に条件分岐によって通過したノードを発火ノードとする。この処理によって、発火および非発火の 2 種類の状態のノードが作成される。その後、注目データ In 0 をモデルのネットワークへ入力する。この際、非発火ノードにおいては、通常のしきい値、発火したノードにおいては発火後しきい値を適用する。この仕組みによって、2 種類のしきい値を単一のノードが扱うことが可能となり、異なるデータ群の影響によって分類結果が変化する分類モデルを表現可能としている。具体的な処理の例として画像およびその特徴量群を扱う場合を述べる。まず、画像を固定サイズのセル状等へ分割する。その後、分割した各セルに対して順に分類処理を行う。この分類処理において、注目セルにおける特徴量群を注目データ群（In 0）、上下左右などの周辺セルにおける特徴量群を参照データ群（In 1, In 2...）とする。注目セルの分類処理において、周辺セルの特徴量群の値を入力し、その値を基に注目セルにおける分類処理のしきい値を変動させる。つまり、周辺セルの特徴量の影響を注目セルの分類結果へ反映させることが可能である。

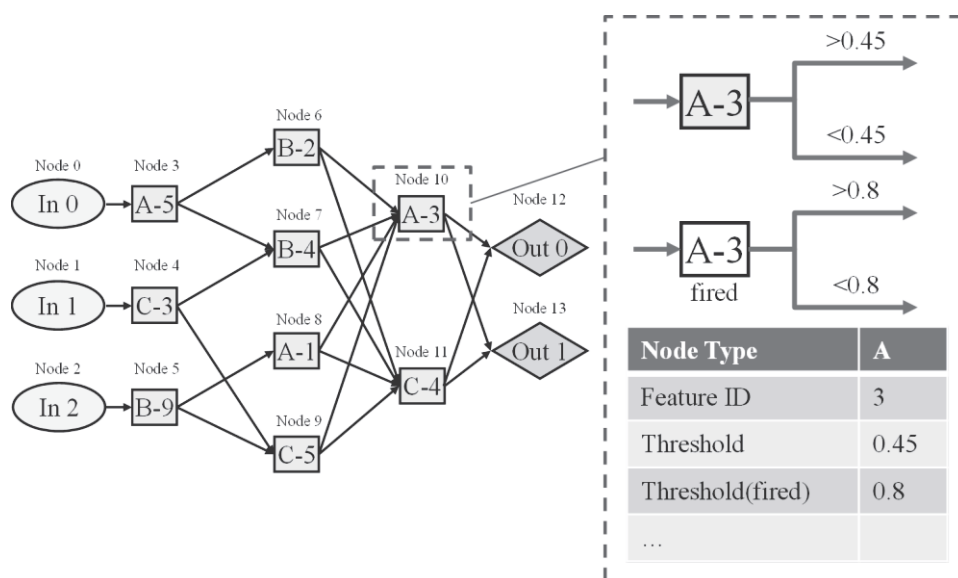


図 5.2 EDEN における異なるデータ群間の相互作用の例

Fig. 5.2 Example of interaction between different data in EDEN

5.2.4 EDEN における遺伝子操作

図 5.3 に EDEN の遺伝子型の例を示す. EDEN の遺伝子型は一次元の数字列である. EDEN では遺伝子型に対して進化計算を行い, その遺伝子型から表現型であるモデルを構築する. 図 5.3 の遺伝子型の一部は図 5.1 の表現型に対応する遺伝子型である. 遺伝子型から表現型への変換は一意に決定される. 遺伝子型における染色体はネットワークを構成する各ノードに対応しており, 各ノードのしきい値 (実数) にあたるパラメータ, 分岐方法 (整数), 分岐先 (整数) を値として保持している. ノードの出力数は, 分岐方法の指定によって決まる. 例えば, ノードの分岐方法が 2 分岐, 3 分岐, 4 分岐まで可能とした場合, 各ノードは 4 つの分岐先および 3 つのしきい値を遺伝子の配列として保持することになる. ここで最大分岐数より少ない分岐数 b が指定された場合を考えると, b 番目以降のしきい値および $b+1$ 番目以降の分岐先は使用しないこととなる. 図 5.3 における中間ノードにおける最大分岐数は 2 であるため, 図中の Node 3 および Node 4 が示すように, 分岐先に関する染色体は 2 つ, しきい値は 1 つのみ保持している. 各ノードの番号はあらかじめ割り振られており, 入力ノード, 中間ノード, 出力ノードの順に番号が大きくなる. フィードフォワード構造の制約を満たすために, 分岐先 (出力先) として接続できるノードは, 接続元のノードが持つ番号より大きい番号を持つノードのみである.

本研究における EDEN の遺伝操作には, GA を用いその世代交代モデルとして Minimal Generation Gap (MGG) [97]を用いた. MGG は, 佐藤らによって提案された GA における世代交代モデルの一つであり, 探索初期における選択圧の抑制することによって局所解に陥り易い初期収束を回避と個体集団内の耐用性を保持すること重視している. このモデルは GA の進化における世代間での個体分布の差の最小化を図ることが望ましいという考えに基づいている. 整数の染色体においては, 一様交叉を用い, 個体の確率選択方法には, ルーレット選択を採用した. また, EDEN の遺伝子型には実数の染色体があるため, その交叉方法においては RCGA [98]における実数値交叉手法の一つである BLX- α 交叉 [71]を用いた.

以下に本研究における世代交代の処理の一例を記述する.

- (a) 個体群 $P(j)$ をランダムに作成し, 全ての個体の適応度を評価する. ($j=0$)
- (b) 個体群 $P(j)$ から非復元抽出により 2 個体 $C_1(j)$ をランダムに選択する.
- (c) $P(j)=P(j)-C_1(j)$ とする.
- (d) $C_1(j)$ より子個体集合 $C_2(j)$ を生成し, 評価する.
- (e) $C_1(j)$ と $C_2(j)$ より最良個体 1 個体と, ルーレット選択によって選択した個体を $P(j)$ に加え $P(j+1)$ とする. ($j=j+1$)

終了条件を満たすまで処理(b)から処理(e)を繰り返し行い, 終了条件を満たした際の最良個

体を学習結果のモデルとして採用する。本研究では、世代数を終了条件として用いた。

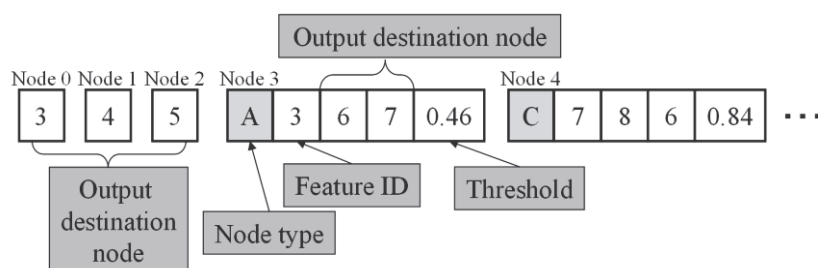


図 5.3 EDEN の遺伝子型の例

Fig. 5.3 Example of a genotype in EDEN

5.2.5 EDEN に使用するノードの拡張

通常の EDEN におけるノードでは単一の特徴量のみ注目して条件分岐を行う。このため、データ間の相互作用を表現することはできない。一方、複数のデータ群を用いた多入力ではデータ群同士の相互作用の表現は可能である。しかし、この相互作用の表現において参照入力とするデータには制約がある。まず、同様な特徴量群を持った複数のデータ群が必要なことである。これは、EDEN においては、注目入力も参照入力も同じネットワークを共有することに起因する。各ノードにおいては、参照する特徴量の番号およびそのしきい値に関する情報が保持されている。つまり、各ノードへ特徴量の数や種類が異なるデータ群が入力された場合に不具合が発生する。具体的には、特徴量の数が異なると参照不可能な番号が発生する。また特徴量の種類が異なるとノードの判定処理において、異なる特徴量を参照するが、しきい値は共有しなければならないという問題が発生する。以上から、通常の EDEN においては、単一のデータ群における特徴量間の相互作用を表現することが難しく、構築可能な構造の自由度が低いと言える。このため、紙の分類実験において複数種類のサンプルを異なる測定器から得られた情報群を基に分類することを考えると、同様な特徴量群を持ったデータ群同士の相互作用ではなく、単一のデータ群における異なる特徴量間の相互作用を表現する必要があるため、EDEN を適用する際にはその構造を変更することが必要である。

そこで本研究の EDEN では、単一のデータ群における特徴量間の相互作用を表現する方法を検討した。相互作用の表現方法においては 2 つの方法を検討した。

1 つ目は EDEN 上に判定用ネットワークとは別に、発火専用のネットワークを構築する方法である。ネットワーク同士の関係は参照入力用のネットワークにおける各ノードに関連ノード番号という新規のパラメータを付与することによって表現する。関連ノード番号は、判定用ネットワークにおける各ノードと参照入力用のネットワークにおける各ノードの接続状態を表現するものである。具体的な処理では、まず、データ群を特徴量の種類によって、注目データ群と参照データ群に分割する。この分割においては各特徴量に‘注目’、‘参

照’といった情報を付与する。次に、参照入力用ネットワークへ参照データ群を入力する。この際に分類までに通過したノードを発火状態とする。発火状態のノードは関連ノード番号に基づいて、判定用ネットワークにおける各ノードへ発火状態を伝達し発火させる。その後、注目データ群を判定用ネットワークへ入力する。ノードを通過する際には通常の EDEN と同様に、非発火、発火によって適用するしきい値を変更する。上記の仕組みによって構成の異なるデータ群同士の影響をしきい値のみで表現可能である。

2つ目の方法は各ノードにおいて複数の特徴量を合成し、その合成値を基に条件分岐を行うノード（図 5.4）の追加である。この方法では、複数の特徴量を参照するノード（複数参照ノード）を新たなノードの種類として追加する。追加に際しては図 5.5 に示すように各ノードにおける遺伝子型に第 2 以降の特徴量 ID および合成時の各特徴量の重みを決定する係数を染色体として追加する。複数参照ノードでは、染色体の情報を基に 2 つの特徴量値を基に新たな値を作成し、その値に対してしきい値による分岐処理を行う。合成する特徴量の最大数はパラメータによって指定する。なお図 5.4 および図 5.5 に示す例は最大合成数を 2 とした場合である。

1 つ目の方法については、しきい値判定の切り替えに重点を置いた方法であり、判定用ネットワークにおいては、通常の EDEN よりシンプルな構造となる利点がある。しかしながら参照入力用ネットワークを用いるため、しきい値が変更される仕組みは複雑となり、判定時の過程は複雑になることが懸念される。一方、2つ目の方法は、参照入力は用いずに単一のデータ群から相互作用を表現する方法である。特徴量群から特定の特徴量を選択し、合成することによって新たなベクトルを持った特徴量を作成することが可能である。異なるデータ群からの参照入力がないためモデルの構造は通常の EDEN のモデルよりシンプルとなる。各ノードにおける処理のシンプルさにおいては、通常の EDEN と比べると複雑となるが、ネットワーク構造を複数組み合わせるモデルと比べるとシンプルである。また、最大合成数を低い値に制限すれば人が理解できる範囲の処理となることが期待できる。例えば、2次元のグラフであれば、画面上に容易にプロット可能であることから、通常の人々の認識能力でその分布特性を認識することが可能である。そこで本研究では、後者の方法を用いて検証を行うこととした。

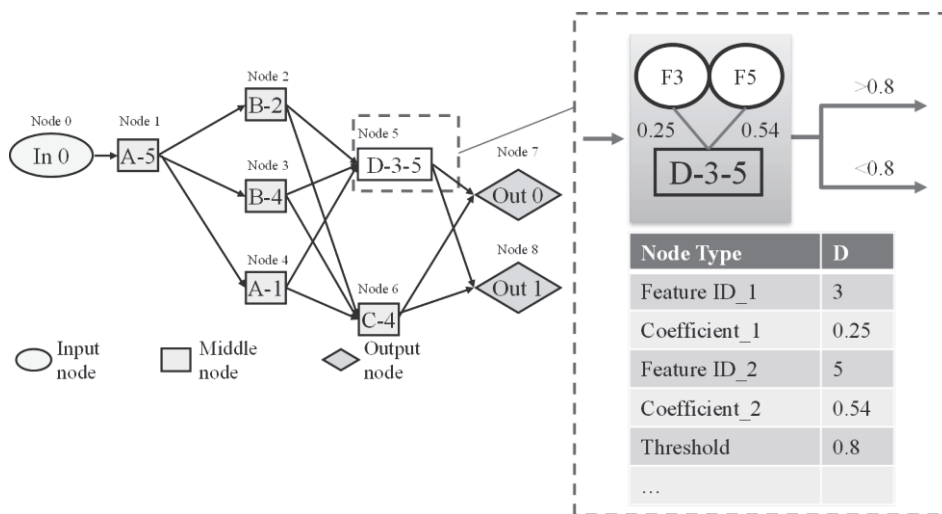


図 5.4 EDEN において拡張したノードの例

Fig. 5.4 Example of extended node in EDEN

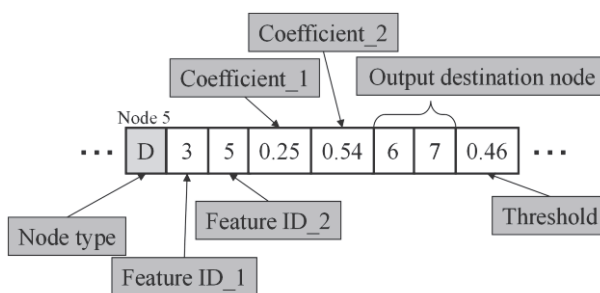


図 5.5 EDEN において拡張したノードの遺伝子型の例

Fig. 5.5 Example of extended genotype in EDEN

5.2.6 ノードを拡張した EDEN の検証

EDEN とノードを拡張した EDEN において、その分類性能やモデルを比較する実験を行った。実験条件としては、表 5.1 に示すような同様の学習パラメータを設定し、対象データには、FS-SVM の検証に用いた Test5 の 5-2 を使用した。学習においては、10 回の試行を行い、最も良い結果が得られた試行時の結果および学習時の最終適応度の平均値について比較を行った。表 5.2 に各手法における最良個体の最終適応度、未知データに対する正答率、モデル内のノード数を示す。最終適応度においては、拡張 EDEN の適応度の値が高く、未知データに対する正答率においても、拡張 EDEN の値が若干高いことが確認できる。また、使用したノードの数は同数であり、モデルの複雑さにおいては同等であった。次に、学習の傾向として、各手法の学習 10 試行における最終適応度の平均値の比較を表 5.3 に示す。適応度の平均値において、拡張 EDEN の値が EDEN より高く、EDEN と比較して、より安定して学習が進んでいることが確認できる。

以上の結果から、本研究において考案した拡張 EDEN は、相互作用の表現およびモデルの構造の自由度の改善において有効であると考えられる。そこで、本研究における EDEN の構造としては、ノードを拡張した考案方式を用いることとした。なお、以降の記載においては、ノードを拡張した EDEN のことを EDEN と記載する。

表 5.1 EDEN および拡張 EDEN の学習パラメータ設定

Table 5.1 Learning parameter of EDEN and extended EDEN

(Parameter for GA)	
Model of iterations	MGG
Population Size	300
Children size	15
Selection method	Tournament selection
Crossover method	Uniform crossover
	BLX- α
Crossover rate	0.6
Rate of uniform crossover	0.1
Mutation rate	0.05
Fitness function	Correct number + (1/ graph edgenumber)
Termination criterion	Generation number
Number of iteration	1,000,000
(Parameter for EDEN)	
Input node number	1
Middle node number	100
Output node number	2
Branch method type	one variable reference composition of multiple reference
Branch number	From 2 to 4
Maximum number of feature composition	2

表 5.2 EDEN と拡張 EDEN の分類性能の比較 (最良個体)

Table 5.2 Comparison of performance of models created by EDEN and extended EDEN
(best individual)

	EDEN	EDEN (extended)
Final Fitness	82.06	90.05
Unknown Correct (%)	0.65	0.67
Number of nodes	9	9

表 5.3 EDEN と拡張 EDEN の平均適応度の比較 (10 試行)

Table 5.3 Comparison of average fitness of EDEN and extended EDEN (10 tests)

	EDEN	EDEN (extended)
Average Fitness	81.96	85.86

5.3 実験

市販の印刷用紙および作製条件の異なる手すきシートをテストシートとし、EDEN によって各テストシートの分類器を構成する実験を行った。

5.3.1 テストシート

テストシートとして、第 4 章で使用したメーカーの異なる PPC 用紙および製造条件の異なる手すきシート（第 4 章 4.3.1 項を参照）を使用した。各テストシートの撮像は光学スキャナ（第 4 章 4.3.2 項を参照）を用いて測定を行った。測定においては用紙の向きを統一し、ワイヤー接触面を光源側に向け、測定を行った。撮像条件は、波長領域を近赤外、輝度分解能 256 階調（8BIT）、空間分解能 400dpi（1 画素=63.5 μ m 角）とした。そして、撮像した測定画像から縦横 200 画素（約 12.5mm 角）の領域を 1 枚の紙片とする紙片画像を手すきシートごとに複数枚作成した。この際、各画像群において紙片の切出し位置が偏らないようにテストシートの全体画像からランダムに切出しを行った。

5.3.2 画像特徴量の算出

各紙片画像において算出した特徴量を以下に記載する。各特徴量の詳細については、第 4 章 4.3.4 項に記載の内容と同様である。

① 階調値の統計量（平均階調，分散等）	17 種
② 2 値画像から得られる統計量（領域数，面積等）	15 種
③ 上下または左右に分割した画像から得られる統計量（共分散，相関等）	6 種
④ GLCM から得られる特徴量（Energy，Contrast 等）	200 種

5.3.3 分類対象

作成する分類器の対象を以下に示す Test 1 から Test 6 となるように設定した。各 Test は 2 分類問題である。詳細については、第 4 章 4.3.6 項に記載の内容と同様である。

- ① Test 1：メーカーの異なる紙の分類
- ② Test 2：原料の異なる手すきシートの分類
- ③ Test 3：紙料の叩解回数の異なる手すきシートの分類
- ④ Test 4：紙料における定着助剤の添加率が異なる手すきシートの分類
- ⑤ Test 5：紙料におけるてん料の転化率が異なる手すきシートの分類
- ⑥ Test 6：シート作成時のワイヤーが異なる手すきシートの分類

5.3.4 EDEN の学習パラメータ

本実験において設定した EDEN の学習に関するパラメータを表 5.4 に示す。最適化手法には GA を採用し、その世代交代モデルに MGG を用いた。選択方式はトーナメント方式、交叉方法には整数部は一様交叉、実数部は BLX- α 交叉法を用いた。適応度関数は以下の (5.1) 式に示すように分類正答数にグラフエッジの逆数を加えたものとした。グラフエッジの逆数は、モデルがシンプルな構造を取るような圧力として加えた。

$$\text{Fitness} = N_c + (1/\text{Edge}) \quad (6.1)$$

N_c は正しく分類された数であり、 Edge は EDEN が構築した分類器モデルにおけるエッジの総数を示す。拡張したノードである複数参照ノードに関するパラメータである最大合成数は人に対する理解のし易さを考慮すると 3 次元以上の特徴量の合成は理解が困難になることを考慮し 2 とした。その他の値は予備実験によって適切な値となるように調整した。

表 5.4 EDEN の学習パラメータ設定

Table 5.4 Learning parameter setting of EDEN

(Parameter for GA)	
Model of iterations	MGG
Population Size	300
Children size	15
Selection method	Tournament selection
Crossover method	Uniform crossover
	BLX- α
Crossover rate	0.6
Rate of uniform crossover	0.1
Mutation rate	0.02
Fitness function	Correct number + (1/ graph edge number)
Termination criterion	Generation number
Number of iteration	1,000,000
(Parameter for EDEN)	
Input node number	1
Middle node number	100
Output node number	2
Branch method type	one variable reference composition of multiple reference
Branch number	From 2 to 4
Maximum number of feature composition	2

5.4 結果

5.4.1 分類結果

各 Test における分類正答率を表 5.5 に示す。正答率は教師画像または未知画像の総枚数に対して正しく分類した画像の枚数を割合で示したものである。正答の定義および正答率の算出は第 4 章 4.4.1 項と同様とした。表 5.5 において、Test 1, Test 2, Test 3, Test 4, Test 6 については高い正答率を示していることが確認できる。しかし、Test 5 については、他の Test と比較して正答率が低いことが確認できる。

表 5.5 EDEN が構築した分類器による各 Test における分類正答率

Table 5.5 Correct answer rate of classifier created by EDEN

Test 1	1-1	1-2	1-3	1-4	1-5
Teacher (%)	100	100	100	100	100
Unknown (%)	100	100	86.7	100	93.3
Test 2	2-1	2-2	2-3		
Teacher (%)	100	100	100		
Unknown (%)	100	100	100		
Test 3	3-1	3-2	3-3		
Teacher (%)	100	100	100		
Unknown (%)	100	94.8	100		
Test 4	4-1	4-2	4-3		
Teacher (%)	100	100	98		
Unknown (%)	100	100	90.6		
Test 5	5-1	5-2	5-3		
Teacher (%)	96.1	97.1	95.1		
Unknown (%)	68.8	52.1	54.2		
Test 6	6-1	6-2	6-3	6-4	
Teacher (%)	100	100	95.6	99.3	
Unknown (%)	100	97.7	81.25	83.6	

5.4.2 構成された分類器

図 5.6 は EDEN によって構築された分類器モデルの一部である。各分類器モデルはそれぞれ、Test 2-1, Test 3-1, Test 4-3, Test 5-2, Test 6-2 において構築されたモデルである。Test 2-1 および Test 3-1 においては単一の画像特徴量のみ, Test 4-3 および Test 6-2 においても比較的シンプルな構造が得られたことが確認できる。Test 5-2 においては, 他の Test と比較すると複雑な構造となっている。また, 本研究で拡張した複数参照ノードが使用されていることが確認できる。

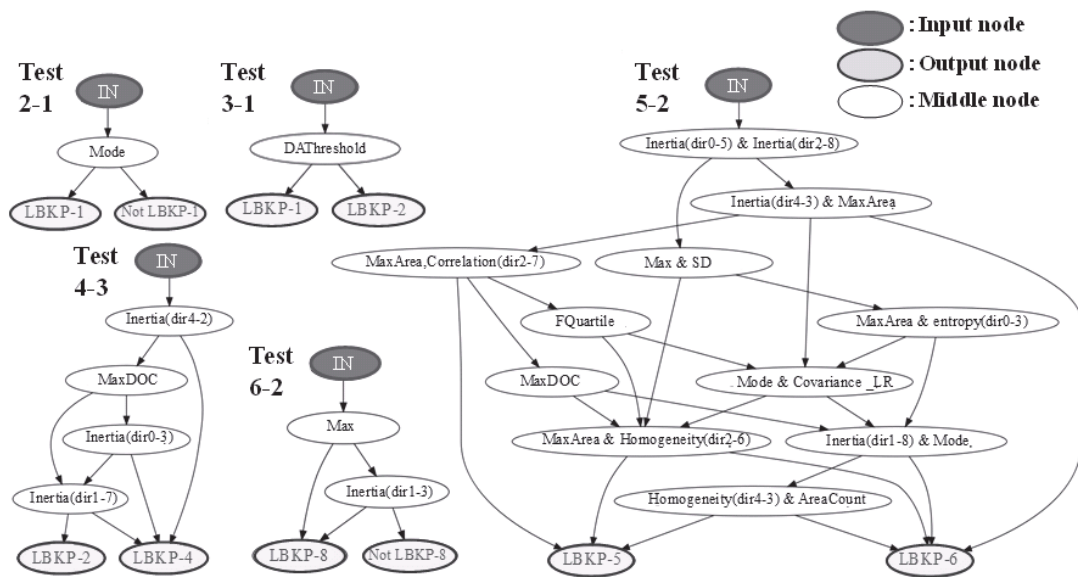


図 5.6 EDEN によって構築された分類器モデルの一例

Fig. 5.6 Example of classifier model created by EDEN

5.5 考察

5.5.1 他の手法との分類正答率の比較

図 5.7 は第 4 章で検証した従来手法である PS による分類結果および特徴量選択型 SVM の分類結果と、EDEN の分類結果を Test ごとに示したものである。グラフにおける値は、各 Test における未知データに対する分類正答率の平均値である。まず従来手法である PS 法は Test 1 および Test 6 において良好な分類結果を得られていることが確認できる。しかし、Test 2, Test 3, Test 4, Test 5 においては低い分類正答率となっている。Test 1 においてはメーカー、Test 6 においてはワイヤーを意図的に変更したサンプルであることから、PS 画像において差が顕著に表れたことが理由として考えられる。一方、機械学習法である特徴量選択型 SVM および EDEN においては Test 5 を除いたすべての Test において高い分類正答率が得られている。EDEN と特徴量選択型 SVM の比較においては、全体的に特徴量選択型 SVM の分類正答率がわずかに高いが、ほぼ同等の分類正答率が得られている。以上から、高い分類正答率を維持しながら分類過程の説明が容易なモデルを構築する点において EDEN は良好な結果を示したことを確認できた。

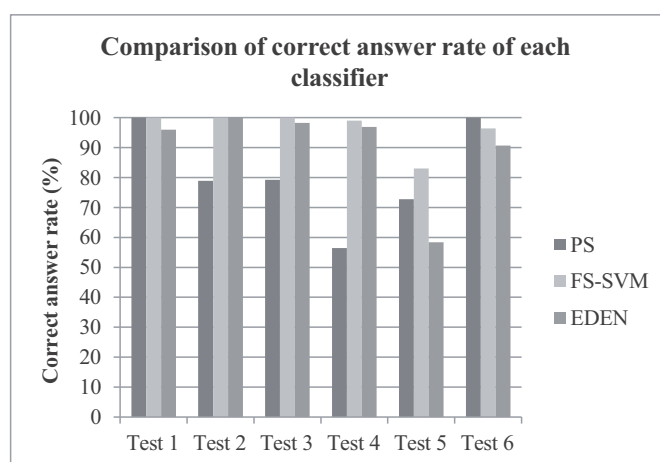


図 5.7 各手法の分類正答率の比較

Fig. 5.7 Comparison of correct answer rate between each method

5.5.2 分類器の構造について

図 5.6 に示した分類器の例を確認すると、Test 2-1 や Test 3-1 など分類が容易な問題においては非常にシンプルな構造が得られている。一方、Test 5-2 のような分類が難しい問題においても、分類器の構造は比較的シンプルで人が理解できる範囲であると言える。Test 5-2 の分類器においては複数参照ノードが使用されていることが確認できるが、その処理は 2 つの

特徴量の重みによる単純な合成処理のみであることから、その分岐の過程は容易である。以上から、複数参照ノードは、複雑な問題のモデルを構築する際に相互作用をシンプルに表現する点において有効であったと考えられる。以上のような EDEN によるモデルにおいて FS-SVM によるモデルとの比較を含め、製紙に関する研究者および製造者と意見交換を行ったところ、以下のような意見が得られた。

- ・ FS-SVM と比べ、各特徴量の関係が確認できる点で有効である。
- ・ 製紙に関する既知の概念以外の情報や特徴量関係に気付くために有用である。
- ・ グラフ構造が複雑になると理解し易いとは言えない。
- ・ 分類過程が可視化されても製造条件などへ直接フィードバックできるわけではない。

意見交換の内容を整理すると、EDEN を用いた手法は分類過程を可視化することにおいて一定の効果はあるが、製造工程における実問題の適用においては、モデルの表現方法および分類性能においてさらなる改善が必要であるとされた。

各モデルおよびその分類性能について、より詳細に確認すると、Test 5-2 における分類正答率においては、教師データに対しては十分な結果が得られているが、未知データに対して分類正答率が低いことが確認できる。この要因としては、第 4 章にも述べたように分類に有効な特徴量が入力データに含まれていないこと、または複数参照ノードの追加によって自由度が向上したことによる過学習の発生などが原因として考えられる。以上から、入力データにおいては、本実験で使用していない他の特徴量の入力を検討する必要があると考えられる。また EDEN においては、適応度関数や複数参照ノードにおける値の合成の方法について、改善が必要であると考えられる。

5.6 まとめ

実験においては EDEN を用いて分類器を構成することによって、分類精度を維持しつつ分類過程が人の目によって確認することが可能な紙の分類器を各 Test に対して構築することができた。また、本研究で拡張したノードである複数参照ノードが複雑な問題の分類器のモデルに使用されていたことから単一のデータ群における特徴量間の相互作用の表現として有効であったと考えられる。分類器モデルにおける複数参照ノードは特徴量間の相互作用の存在の確認、ノード数の削減による分類器モデルのわかり易さの点において有効であった。以上から、EDEN による紙の分類モデルの構築においては、分類精度の向上および人に理解のし易い分類過程を得るという課題において一定の成果が得られたと言える。しかしながら、適応したデータによっては、未知データに対する分類正答率が低いという結果も見られ、一部においては FS-SVM と比較して分類性能が低い結果もあった。この点においては、分類過程の可視化という成果があった一方で、分類性能の向上は最も重要な課題であることから改善する必要がある。そして、今回設定した適応データである各 Test は、紙の特性として大きく変動させたサンプルが主体であった。製造工程における実問題においては、分類精度のさらなる向上が必要とされることが想定されるため、入力する特徴量および分類器の構造を改善し、より複雑な問題における検証および検証が必要である。

第 6 章 Cartesian Genetic Programming を用いた紙の特性予測器の自動構築

6.1 はじめに

第 4 章および第 5 章では、機械学習によって製造された紙をその特性によって分類する手法について検討を行い一定の成果が得られた。そこで次のステップとして製造前および製造時の情報から紙の特性を予測する手法について検討を行った。紙の製造において材料や製造環境から得られる情報を活用することは品質を向上させるための重要な手段であり、従来からパルプ紙料の特性や製造条件から紙の特性の予測を行う研究が行われている（第 2 章および第 3 章を参照）。紙の製造においては、特にパルプ紙料と紙の特性の関係が重要とされている [99]。上記の研究においては重回帰分析（Multiple Linear Regression; MLR）や神経回路網（Neural Network; NN）を用いた解析が行われており、予測精度の向上について、一定の成果が得られている。統計的手法である MLR は、結果と要因の関係を単純な予測モデル（結果と要因の定量的な関係を表す式）で表現する。このため、結果と要因の関係の説明が容易である。しかし、MLR は複雑な問題において、高い予測精度と汎化性能を向上させることが難しい。一方、NN は、機械学習によって予測モデルや分類器を構築する手法で、複雑な問題に柔軟に対応可能である。しかし、NN によって構築された予測モデルは複雑な構造となる傾向があるため、予測結果と要因の関係の説明が困難である。そこで本研究では、Cartesian Genetic Programming (CGP) [100]を基にモデルを構築する方法を検討した。CGP は MLR と比べ柔軟なモデルを構築することが可能である、また、単純なグラフ構造によって予測モデルを構築することから NN と比較して、結果と要因の関係の説明が容易になることが期待できる。本研究では、CGP を基にシンプルな構造で予測精度の高いモデルの構築を目指した。実験では、紙料中の繊維の特徴から、紙の光学特性を予測するモデルを CGP, MLR, NN を用いて構築し、その性能について評価した。

G	: 遺伝子型
n_i	: 入力数
n_o	: プログラムの出力数
n_n	: ノードの入力数
F	: ノードの関数群
n_f	: 関数の数
n_r	: ノード配列における列の数
n_c	: ノード配列における行の数
l	: レベルバックパラメータ (ノードの接続制限に関する値, ノードが現在の位置から何行前の行に属するノードと接続可能であることを示すパラメータ)

CGP の遺伝子型における整数は特定のノードの処理の内容を示す遺伝子にあたる。ノードの処理を説明する 1 組の整数群における整数の数は、入力数 n_n と関数を指定する番号で構成されるため、 (n_n+1) 個となる。つまり、遺伝子型における遺伝子はノードの総数である $(n_r * n_c)$ 個に出力数 n_o 個を含めたものとなる。例えば、パラメータを以下のように設定する。

n_i	: 2
n_o	: 1
n_n	: 2
n_f	: 4
n_r	: 4
n_c	: 4
l	: 4

そして、入力 0 および入力 1 をそれぞれ $\{1.0, X\}$ とし、関数群 F を $\{+, -, *, div\}$ とし、その番号に 0,1,2,3 をそれぞれ割り当てる。以上の条件の CGP 遺伝子型の例を以下に示す。

(1.0 ; X ;
 1, 0, 3 ; 1, 1, 3 ; 1, 1, 2 ; 0, 0, 2;
 4, 4, 2 ; 2, 1, 3 ; 3, 4, 1 ; 1, 4, 0;
 6, 3, 2 ; 1, 8, 2 ; 0, 8, 2 ; 8, 6, 2;
 6, 11, 2 ; 11, 11, 2; 5, 3, 3 ; 13, 13, 3;
 15)

各ノードの番号はそれぞれ 0 から 17 の整数で表されている。例えば、上記遺伝子型において番号が 3 のノードは(1,1,3)を示している。各ノードのうち、出力に関係するノードは下線で示されている。左端の 2 つの値、1.0 と X は入力値である。右端の 15 は出力元となるノードの番号を示している。下線が引かれた各ノードを以下に示す。

Node 3 : $X/X = 1.0$

Node 4 : $X*X = X^2$

Node 8 : $1.0-X^2$

Node 11 : $X*(1.0-X^2) = X-X^3$

Node 15 : $(X-X^3)*(X-X^3) = X^6-2X^4+X^2$

したがって、この遺伝子は数式 ' $X^6-2X^4+X^2$ ' を表している。図 6.1 はに上記の遺伝子型の表現型を示したものである。

6.2.2 CGP の構造の検討および拡張

本研究では基本的な CGP を基に、よりシンプルかつ自由度の高いモデルを構築することを目的に、CGP のパラメータを調整すると同時に、遺伝子型の構造を変更することによって、通常の CGP では実行が難しい処理を表現可能にした。以下にその特徴を述べる。

① パラメータの検討

CGP のノード配列において同じ行に属するノードは互いに接続することはできない。CGP における最も自由なネットワークは $n_r=l$ かつ $l=n_c$ のときである。つまり、どのノードも自分より大きいノード番号を持つノードとであれば、自由に接続可能な状態である。遺伝子型の長さはパラメータによって $n_r n_c (n_m + 1) + n_o$ に固定される。しかし、このパラメータは CGP モデルの最大サイズを決定しているだけのものであり、実際のモデルのサイズは 0 から最大サイズの間であれば自由に構築可能である（例えば、図 6.1 の例では 5 つのノードのみでモデルを構成している。）。よって本研究では、ノード配列における列の数は n_r は 1、レベルバックパラメータ l は、ノードの行数と同じとした。

② 遺伝子型への自由変数の追加

予測モデルを構成する際に、しきい値を用いた条件判断処理やゲイン・オフセット処理はよく利用される処理である。これらの処理は、特定の値に関連する情報を基に値を決定し実行される。しかしながら、参照可能な情報群に必ずしも特定の値に関連する情報が含まれているとは限らない。この場合、統計的なデータまたは経験的に値を作成する必要がある。通常の CGP の遺伝子型は整数のみの配列で構成されたため、ある値に対して、しきい値を用いた条件判断処理やゲイン・オフセット処理など実数を使用する処理

が必要とされる場合、特定の実数値を作成する構造をモデル内に構築する必要がある。この構造は、入力値を参照するノード群を必要とする。また、この際の入力値は、対象とする特定の値との関連性は必要とされない。つまり、不要な特徴量を選択する可能性を発生させると同時に、複雑なモデルを構築する要因となる。そこで、本研究における CGP では、モデルの自由度を向上させるために遺伝子型に自由変数を加えた。これによって各ノードが独立した実数を持つことになり、他の入力値の参照や、各ノードの経路の繋がりを必要することなく、上記のような特定の処理を学習によって獲得とすることが可能となり、モデルをよりシンプルに表現することが期待できる。変更後の個体を表すモデルは以下の (6.2) ように定義される。

$$\{G, R, n_i, n_o, n_n, F, n_f, n_r, n_c, l\}$$

上記の定義における R は自由変数を意味し、その値は実数である。自由変数 R は学習過程において、突然変異等によって指定された値の範囲内でランダムに変更される。実数の可変範囲はパラメータで指定する。図 6.2 に本研究で拡張した CGP の遺伝子型および表現型の例を示す。

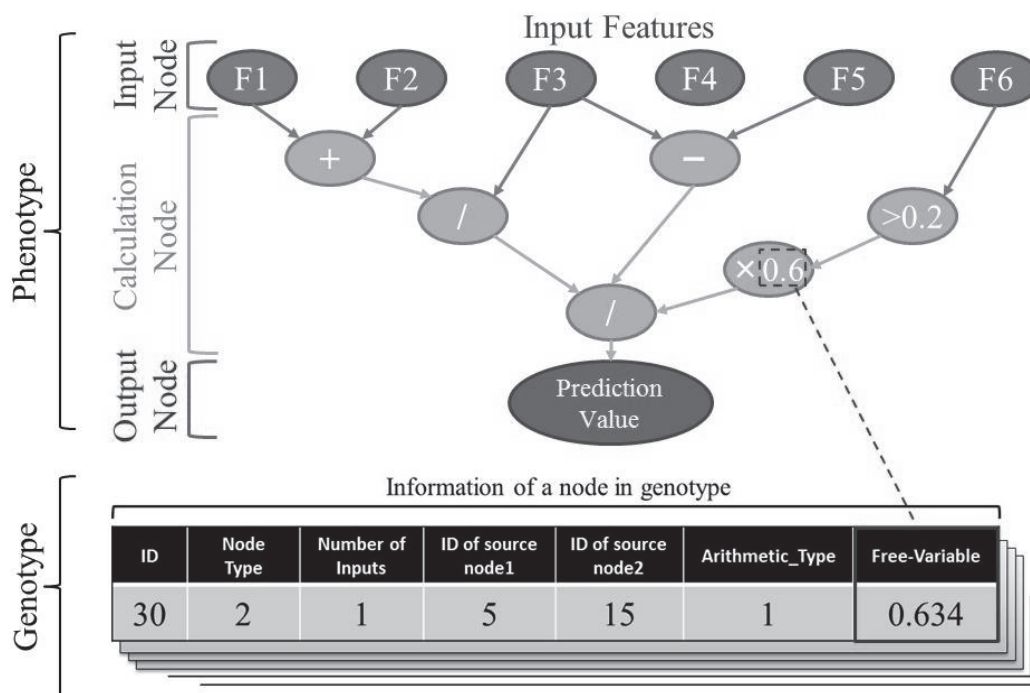


図 6.2 本研究で拡張した CGP の遺伝子型と表現型

Fig. 6.2 CGP phenotype and genotype in this study

6.2.3 拡張 CGP の検証

CGP と拡張 CGP において、その予測性能を比較する実験を行った。実験条件としては、表 6.3 に示す学習パラメータを設定し、対象データには、紙料中の繊維特徴量と同紙料から作製した手すきシート（詳細については 6.3 節に記述する。）を使用した。学習においては、10 回の試行を行い、最も良い結果が得られた試行時の結果について比較を行った。比較の指標としては、実測値と予測値間の R^2 値を採用した。図 6.3 に CGP と拡張 CGP の予測性能の比較結果を示す。図 6.3 から実測値と予測値間の R^2 値において、拡張 CGP が CGP より高い値を示していることが確認できる。また、表 6.1 に各手法によって作成された予測モデルにおいて使用されているノード数および特徴量数を示す。表 6.1 から CGP と拡張 CGP のノード数および特徴量数に大きな差がないことが確認できる。

以上の結果から、本研究において考案した拡張 CGP は、使用ノードおよび特徴量数の増加によるモデルの複雑化の抑制とモデルの自由度の改善において有効性を示した。そこで、本研究における CGP の構造としては、拡張 CGP を用いることとした。なお、以降の記載においては、拡張 CGP のことを CGP と記載する。

図 6.3 CGP と拡張 CGP の予測性能の比較

Table 6.3 Comparison of prediction accuracy between CGP and extended CGP

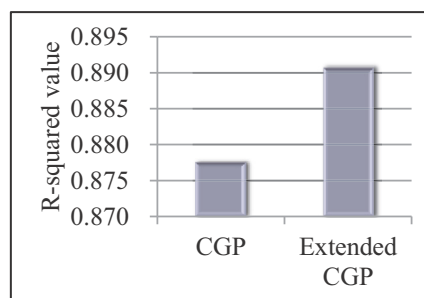


表 6.1 CGP および拡張 CGP のモデルにおいて使用されたノード数および特徴量数

Table 6.1 Number of used nodes and used features of the model created by each method

	CGP	Extended CGP
Node_num	40.5	43.5
Feature_num	13	13.5

6.3 実験

6.3.1 テストサンプル

テストサンプルとして、表 6.1 に示す 9 種類のパルプ紙料を作製した。各サンプルはメーカー、裁刻の有無、原料の混合割合などが異なる。原料には Cotton linter pulp, abaca pulp, SBKP, HBKP を用いた。なお、S08 と S09 は他のサンプルを混合して作成したサンプルである。S08 とは S03 と S04, S09 は S06 と S07 を同じ比率で構成したサンプルである。各サンプルのシートの作製は、手すきによって行い、各サンプル 10 枚ずつ作製した。その後、手すきシート 1 枚当たり 25 の紙片をテストシートとして採取した(各サンプル 250 枚,のテストシート)。手すきシートの作製に当たっては ISO 基準 (ISO 5269-1) に則って作業を行った。なお、表記以外の作製条件は同様となるように調節した。

表 6.2 繊維特徴量のリスト

Table 6.2 Pulp suspension employed for test samples

	Pulp	Lot	Manufacturer	Shredding
S01	Cotton	L1	C	No
S02	Cotton	L2	C	No
S03	SBKP	L3	C	No
S04	HBKP	L4	C	No
S05	Abaca	L5	A	Yes
S06	Abaca	L5	B	No
S07	Abaca	L5	B	Yes
S08	S03 + S04	—	—	—
S09	S06 + S07	—	—	—



図 6.4 作製した手すきシートの一例

Fig. 6.4 Example of hand sheets

6.3.2 繊維特徴量

本実験では、予測するための情報として、各パルプ紙料中の繊維の特徴量を用いた。各特徴量は、図 6.4 に示す繊維特徴量測定器（Metso Automation 社製 kajaani fiberLAB）を用いて測定した。各測定は、ISO16065-1（2001）に準拠して実施した。実験では、fiberLAB から出力される 47 種類の特徴量を用いた（表 6.2）。繊維特徴量測定器では、レーザ検出器と高分解能カメラによって図 6.5 のように繊維一本一本の画像を撮像する。その後、表 6.2 に示す各繊維の特徴量を画像から算出する。本実験における各特徴量は、パルプ紙料中の繊維の特徴（長さ、幅など）やベッセルの特徴（大きさ、個数など）の測定値の平均値を用いた。測定した繊維の本数はパルプ紙料ごとに約 3000～8000 本で、ベッセルは約 450 個である。表 6.2 の括弧内の各記号は算出方法の違いを表す。proj は投影長さを示し、cont は中心線長を測定したものである。n は数平均、l は長さ加重平均、w は質量加重平均を表している。上記のように各特徴量には、類似した算出方法で求められている値もあり、値同士の相関が高い特徴量も含まれる。本実験では、各手法の次元削減の能力を比較するため、人の手による特徴量の選択は行わず、すべての特徴量を情報として利用した。

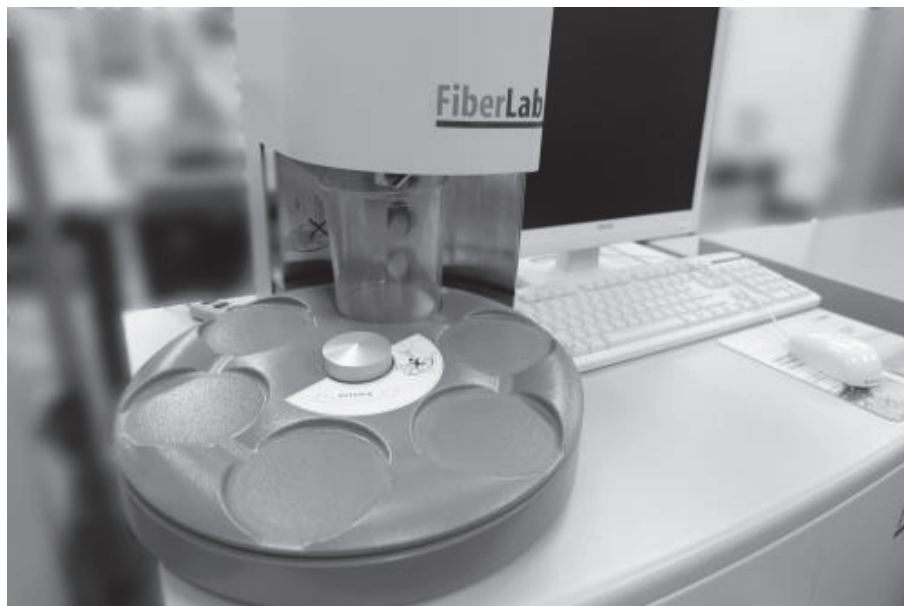


図 6.5 繊維特徴量測定機（Metso Automation 社製 kajaani Fiber Lab）

Fig. 6.5 Fiber feature measuring device (kajaani Fiber Lab by Metso Automation)

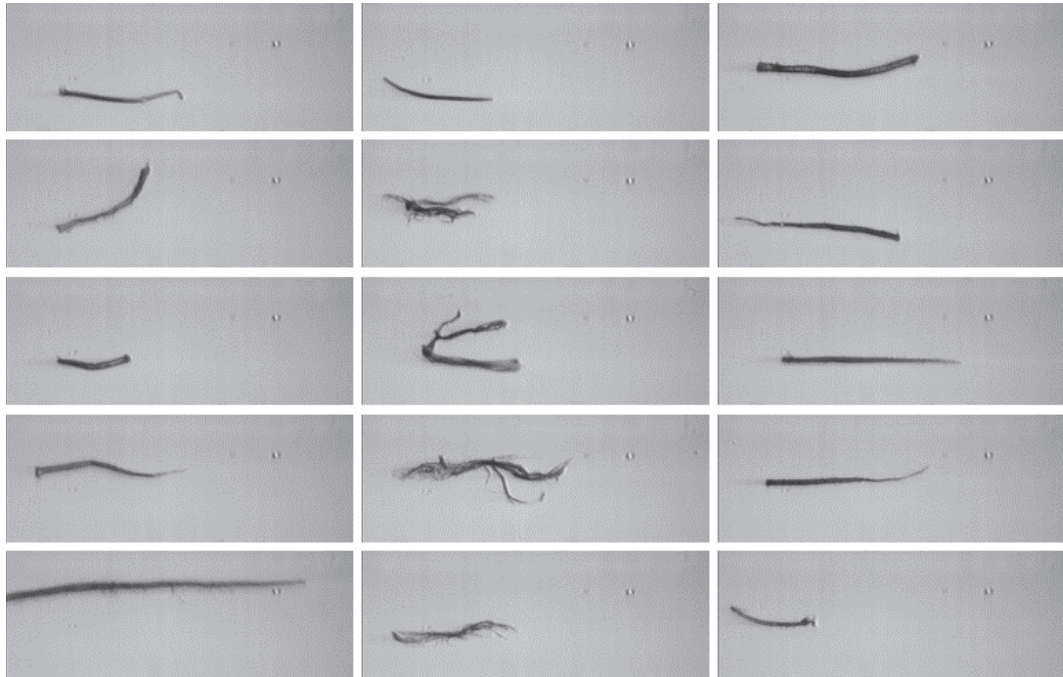


図 6.6 繊維特徴量測定器によって撮像した繊維画像の一例

Fig. 6.6 Examples of fiber images captured by Fiber Lab

表 6.3 繊維特徴量測定器が出力する繊維特徴量

Table 6.3 Fiber feature amounts outputted by Fiber Lab

ID	Name	ID	Name	ID	Name
F01	Length (proj) (n) [mm]	F17	Width (n) [μm]	F33	Vesselwid (l) [mm]
F02	Length (proj) (l) [mm]	F18	Width (l) [μm]	F34	Vesselwid (w) [mm]
F03	Length (proj) (w) [mm]	F19	Width (w) [μm]	F35	Vessel [1/1,000 fibres]
F04	Length (cont) (n) [mm]	F20	Wall (n) [μm]	F36	Vessel [1/m]
F05	Length (cont) (l) [mm]	F21	Wall (l) [μm]	F37	Vessel [1/mg]
F06	Length (cont) (w) [mm]	F22	Wall (w) [μm]	F38	Vessel [$\text{mm}^2/1,000$ fibres]
F07	Fines (proj) (n) [%]	F23	Curl (n) [%]	F39	Vessel [mm^2/m]
F08	Fines (proj) (l) [%]	F24	Curl (l) [%]	F40	Vessel [mm^2/mg]
F09	Fines (cont) (n) [%]	F25	Curl (w) [%]	F41	Fibrillation [%]
F10	Fines (cont) (l) [%]	F26	Kink (n) [1/m]	F42	Csa (n) [μm^2]
F11	ISOL (proj) (n) [mm]	F27	Kink (l) [1/m]	F43	Csa (l) [μm^2]
F12	ISOL (proj) (l) [mm]	F28	Kink (w) [1/m]	F44	Csa (w) [μm^2]
F13	ISOL (proj) (w) [mm]	F29	Vessellen (n) [mm]	F45	Volume Index (n) [$10^6 \mu\text{m}^3$]
F14	ISOL (cont) (n) [mm]	F30	Vessellen (l) [mm]	F46	Volume Index (l) [$10^6 \mu\text{m}^3$]
F15	ISOL (cont) (l) [mm]	F31	Vessellen (w) [mm]	F47	Volume Index (w) [$10^6 \mu\text{m}^3$]
F16	ISOL (cont) (w) [mm]	F32	Vesselwid (n) [mm]		

6.3.3 紙の光学特性

本実験では、紙の光学特性を予測の対象とした。紙の光学特性には、手すきシートの画像（図 6.6）から算出した画像特徴量を用いた。手すきシートの画像は測定装置にはアイメジャー社製の赤外透過型撮像装置（IR-6000，第 4 章，図 4.3 参照）を用い、解像度 400dpi で撮像した。画像特徴量には、画素値の統計量とテクスチャ特徴を用いた。画素値の統計量には Mean と標準偏差（Stdev）を用いた。そして、テクスチャ特徴量には Gray Level Co-occurrence Matrix（GLCM）から算出可能な Energy, Inertia を用いた。ここでテクスチャ特徴量とは、画素値の周期性や振幅の大きさを表した数値である。図 6.6 に記載されている各値は各テストサンプルにおける各光学特性値の平均値である。なお、図 6.6 に記載の画像は本紙面上での視認性を考慮し、画像処理によって階調を補正した画像である。

	S01	S02	S03	S04	S05
Mean	95.9741	105.1606	101.5933	100.6826	100.6826
Stdev	4.0553	4.5143	2.7591	4.1345	4.1345
Energy	0.0091	0.0075	0.0136	0.0064	0.0064
Inertia	5.5016	6.4966	6.6026	13.9087	13.9087

	S06	S07	S08	S09
Mean	86.7531	100.1320	92.7637	98.9890
Stdev	2.2362	3.2495	2.9069	4.0722
Energy	0.0184	0.0097	0.0116	0.0073
Inertia	5.8304	9.4760	8.5186	9.8127

図 6.7 各テストシートの画像の一例

Fig. 6.7 Images of each test sheet

6.3.4 比較手法

多変量解析において一般的に用いられる統計手法である線形重回帰分析，そして機械学習を用いた予測手法において適用されることの多い NN，これら 2 手法を比較手法として選択した。各手法に関する概要および本研究における条件の設定を以下に記載する。

① 線形重回帰分析（Multiple Liner Regression; MLR）

回帰分析（Regression analysis）は、データの解析や予測において、統計的手法の中で最も広く応用されている解析方法である。この手法は回帰方程式と呼ばれる変数間の説明の関係を定量的に表す式を求める。複数の変数間の関係を解析する手法を重回帰分析

(Multiple Regression Analysis; MRA) と呼び、その変数間の関係を線形関数で表すものを線形重回帰分析 (Multiple Liner Regression; MLR) と呼ぶ。また、本研究では、予測モデルの汎化性を向上させるための特徴量の次元削減手法として、Akaike Information Criterion (AIC) [103]を用いた。

② 神経回路網 (Neural Network; NN)

本手法の概要については第 3 章 3.2 節に記載してある通りである。本実験では、NN の構造として単純な階層型 feedforward ネットワークを用いた。階層は 3 層とし、ネットワークは 3 つの Layer , “Input Layer”, “Hidden layer”, “Output Layer”で構成される。また、学習方法には誤差逆伝搬法を用いた。

6.3.5 機械学習に関するパラメータおよび条件

本項では、CGP, NN に関するパラメータおよび条件について記載する。表 6.3 は各手法において設定したパラメータを示している。表 6.3 の各値は予備実験の結果を基に調整した。NN, CGP 共に 10 回の Test を行い、その内最良の結果を示したものを実験結果として採用した。各手法に関するより詳細な説明を以下に記載する。

① NN

すべての入力層のニューロンはすべての隠れ層のニューロンと接続されており、すべての隠れ層のニューロンはすべての出力層のニューロンと接続されている。入力層からは 47 種類の特徴量が入力され、隠れ層は 30 に設定した。従って、入力層と隠れ層の間には 1410 個の接続があり、隠れ層と出力層の間には 30 個の接続がある。ユニット関数にはシグモイド関数を用いた。各接続の重みは BP 法 によって調整した。BP 法では、ネットワークが出力した値 y_k と教師データの値 t_k を Table 4 に示される Error Function によって比較し、その誤差を基に修正量を各接続に伝えることによって、その重みを修正する。本実験では、修正の強さである Learning rate を 0.01, 1 回の学習での繰り返し回数である Training time を 1000 に設定した。

② CGP

CGP のモデル構築における最適化処理には ES を用いた。終了条件は世代数とし表 6.3 の記載した値に設定した。本実験における ES における最適化処理には、 $(1+\lambda)$ ES 法を採用した。本研究における $(1+\lambda)$ ES 法における最適化処理の過程を以下に示す。

- (a) 指定サイズの個体群をランダムに作成し、全ての個体の適応度を評価する。
- (b) 個体群から 1 個の個体をランダムに選択する。
- (c) 選択された個体に突然変異操作を行った新たな個体を λ 個生成する。

- (d) 新たな個体の適応度を評価する.
- (e) 選択された個体と新たな個体のうち, 最も適応度の値の低い 1 個の個体を選択し個体群に戻す. 残りの個体は削除する.

最適化処理において, 処理(a)~処理(e)は, 終了条件を満たすまで繰り返される. そして終了条件は満たしたのち, その時点における最良の個体を結果として採用する. 本実験における終了条件は表 6.3 に示す世代数とし, λ は 4 に設定した. 個体群のサイズは 100 とし, 適応度関数は表 6.3 に示す式を用いた. この関数では, CGP モデルの出力値 y_k と教師データの値 t_k の差を求める. 学習プロセスにおける目標はこの値を最小化することである.

次に本実験において使用した計算ノードについて述べる. 計算ノードのリストを表 6.4 に示す. 各計算ノードは, 1 入力または 2 入力の 2 種類に大別される. 1 入力計算ノードは, 1 つの入力値, 2 入力ノードは, 2 つの入力値を基に計算処理を行い, その結果をノードの出力値とする. 表における Input 1 および Input 2 の表記はそれぞれ第 1 入力と第 2 入力を表している (6.4.4 節, 図 6.10 参照).

表 6.4 各機械学習手法のパラメータ

Table 6.4 Machine learning parameters

NN		CGP	
Training time	1,000	Number of iterations	5,000,000
Type	3-layer	Gene manipulation	(1+4)ES
Learning method	Back propagation	Population size	100
Hidden unit	30	Mutation rate	0.05
Learning rate	0.01	Input nodes	47
Momentum	0.2	Calculation nodes	10,000
Unit function	Sigmoid function	Range of free-variables	± 1.0
Error function	$E_n = \frac{1}{2} \sum_{k=1}^K (y_k - t_k)^2$	Max input	2
		Node grid columns	1
		Node grid rows	10,047
		Level back	10,047
		Fitness function	$Fitness = -\frac{1}{2} \sum_{k=1}^K (y_k - t_k)^2$

表 6.5 CGP に使用した計算ノード

Table 6.5 CGP calculation node types

Two input node		One input node	
Sign	Output	Sign	Output
+	Input 1 + Input 2	Reverse	Inversion of positive/negative
-	Input 1 - Input 2	Exp	Exp function
*	Input 1 × Input 2	Log	Log function
/	Input 1 / Input 2	Pow	Second power of Input 1
Greater	0 (Input 1 < input 2) 1 (Input 1 > input 2)	Sqrt	Square root of Input 1
Less	0 (Input 1 > Input 2) 1 (Input 1 < Input 2)	Abs	Absolute value of Input 1
Equal	0 (Input 1 and Input 2 are not equal.) 1 (Input 1 and Input 2 are equal.)	Thresh	0 (Input 1 is between 0 and 1.) 1 (other)
		1	1 (fixed value)
		0	0 (fixed value)
		0.1	0.1 (fixed value)
		* (value)	Multiplication of Input 1 and free variable
		+ (value)	Addition of Input 1 and free variable
		> (value)	0 (Input 1 < free variable) 1 (Input 1 > free variable)
		< (value)	0 (Input 1 > free variable) 1 (Input 1 < free variable)
		Piecewise	Piecewise function Input 1 (0 < Input 1 and Input 1 < 1) 1 (other)
		Sigmoid	Sigmoid function
		Normal	Normal distribution function

6.3.6 予測モデルの検証方法

各手法によって作成された予測モデルの予想性能の評価においては S fold 交差検定を実施した。S fold 交差検定とは、予測手法の妥当性を検証する方法である。この手法では、データをランダムに S 個のデータに分割する。S - 1 個の分割データを学習データ（教師データ）として使用し、残り 1 つの分割データを、予測モデルをテストするためのテストデータ（未知データ）として使用する。そして、各分割データがテストデータとなるように組み合わせを変更し、計 S 回の検証を行う。最終的には、S 回の平均を検証結果として用いる。本実験では set1 から set5 までの 5 つの dataset を設定した (S=5)。各 dataset に含まれるサンプルの数は合計で 2250 である（教師データ 1800、未知データ 450）。

6.4 結果

6.4.1 CGP の学習における適応度の推移

図 6.7 は CGP による予測モデル学習時における最良固体の適応度 (fitness) の推移を表したグラフである。各適応度は表 6.3 に示した適応度関数 (Fitness function) によって算出される。各グラフは 4 つの異なる光学特性に対する学習時の様子を示している。グラフの縦軸は適応度、そして横軸は世代数を表している。グラフから各世代を重ねるごとに適応度が上昇していることが確認できる。つまり、予測モデルの最適化が進んでいることを示している。また、適応度は 100 万世代目までに大幅に上昇し、それ以降の世代では上昇率が低い。これは、前半の世代で、予測モデルの主要な構造が得られていることを示している。

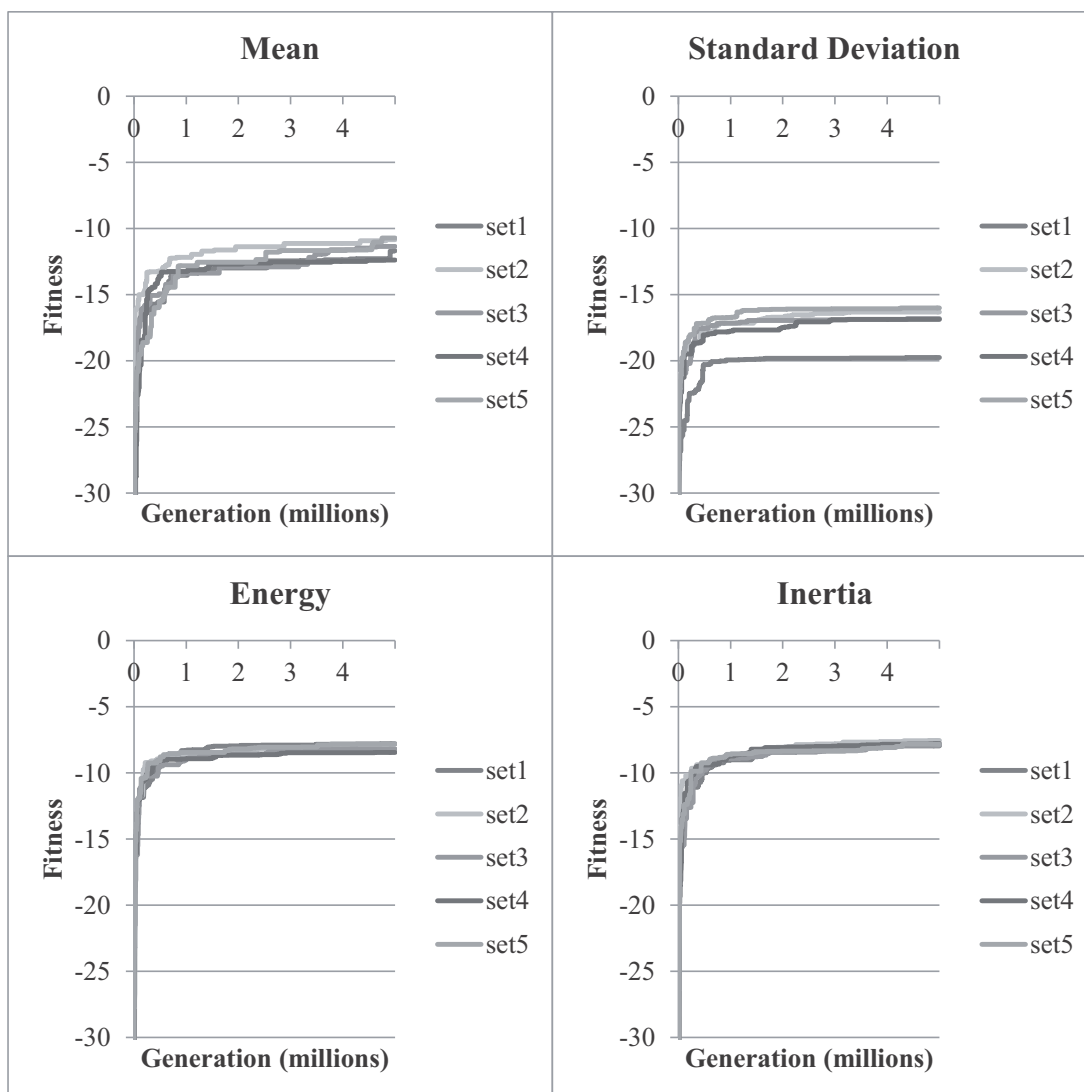


図 6.8 CGP における適応度の推移

Fig. 6.8 CGP fitness transition

6.4.2 予測精度の比較

各手法の予測精度を図 6.8 に示す。縦軸は予測値と実測値間の R-squared value を示し、横軸は各手法を示している。予測値とは、パルプ紙料中の繊維の特徴量から算出した値であり、実測値とは同パルプ紙料から作製したテストシートの光学特性である。グラフ左側は教師データを適用した際の結果、グラフ右側は未知データを適用した際の結果である。各値はデータセットの set1 から set5 の平均値である。まず、教師データに対する R-squared value は、各手法とも高い値を示している。つまり、各手法が教師データから判断基準を獲得し、予測モデルを構築したことを示している。次に、グラフ右側は未知データへの適応結果を示している。グラフでは全ての予測対象 (Mean, Standard Deviation, Energy and Inertia) に対して CGP が最も高い R-squared value を示している。この結果は、CGP によって構築された予測モデルが、他の 2 手法より高い精度の予測値を算出したことを示している。

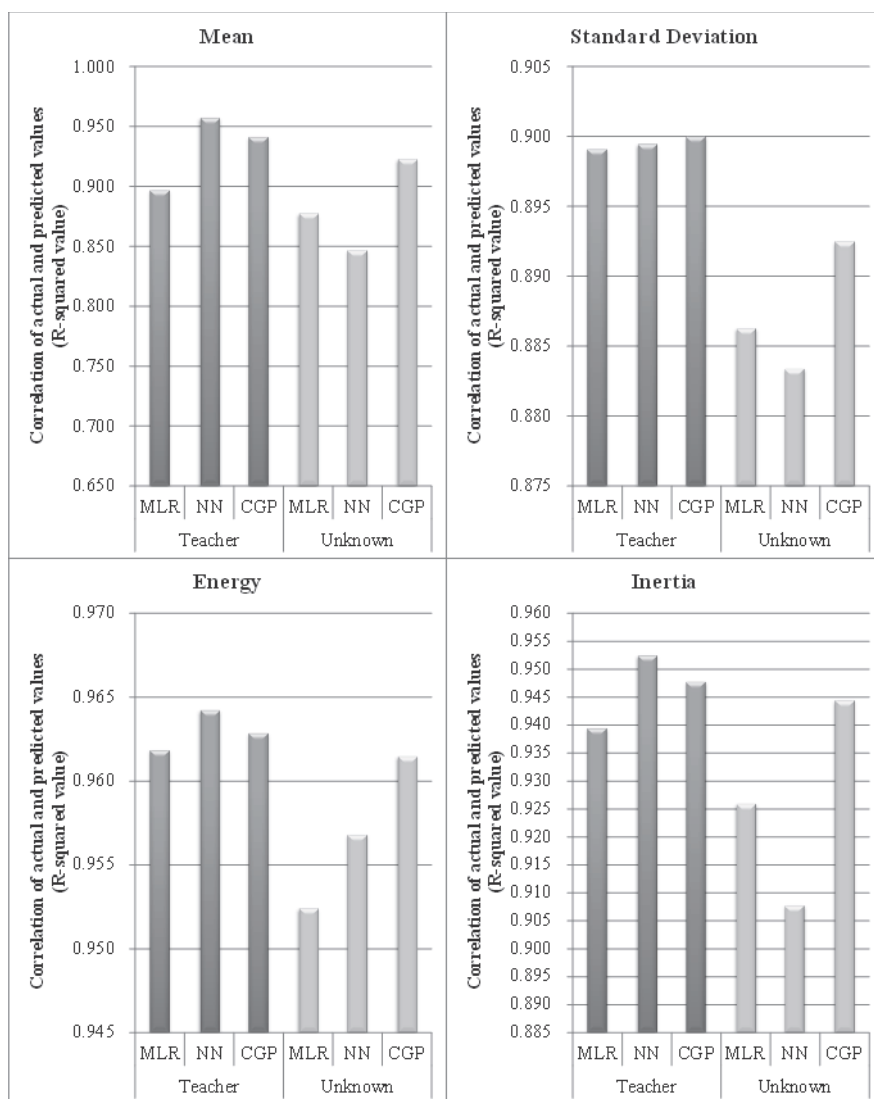


図 6.9 各手法によって構築されたモデルの予測精度の比較

Fig. 6.9 Comparison of prediction accuracy levels obtained with each method

6.4.3 次元数の削減の比較

図 6.9 は各予測モデルの構築において MLR と CGP がそれぞれ選択した特徴量の数を示している. このグラフでは CGP によって構築された予測モデルの使用特徴量数が MLR によって構築されたモデルより少ないことを示している. 図 6.8 が示すように予測性能において他の 2 手法より良好な結果が得られていることを合わせると, 今回の実験においては CGP による次元削減が有効に働いたことを示している.

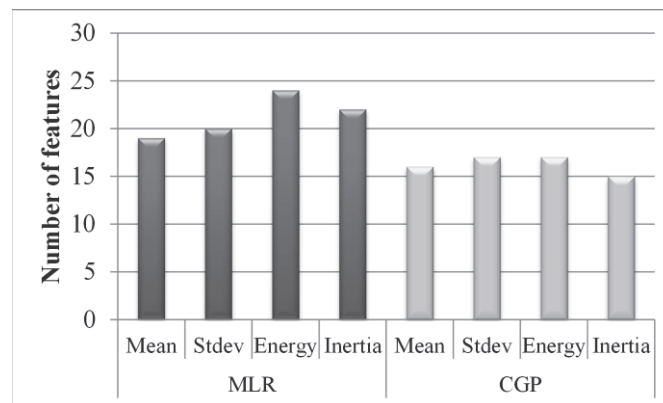


図 6.10 予測モデルに使用された特徴量数の比較
Fig. 6.10 Comparison of numbers of employed features

6.4.4 CGP が構築した予測モデルの一例

本実験において CGP が構築した予測モデルの構造の一例を図 6.10 に示す. この図は予測値の算出過程を示している. このモデルは Mean の予測モデルである. このモデルが予測に使用した特徴量数は 15, 計算ノード数は 32 である.

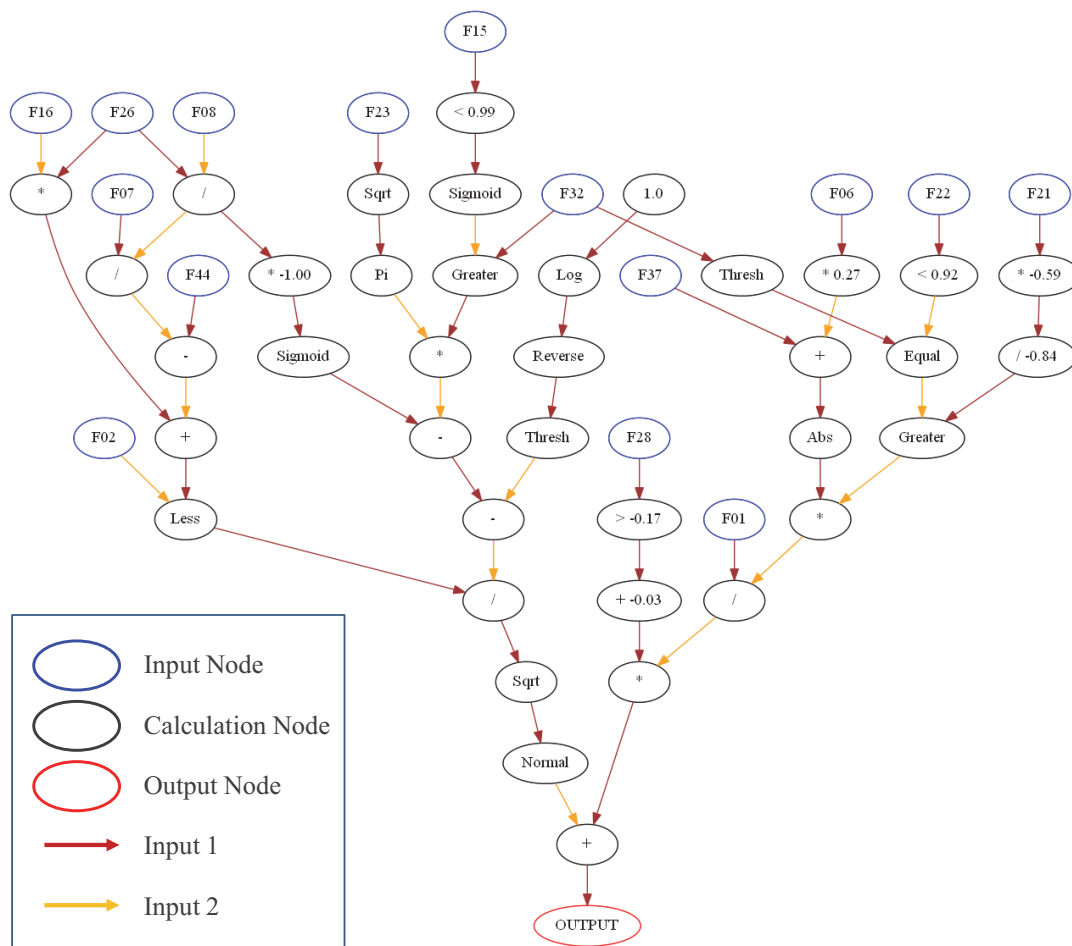


図 6.11 CGP によって構成されたモデルの例
Fig. 6.11 Example of prediction model created using CGP

6.5 考 察

6.5.1 各モデルの予測性能

本実験では CGP の未知データに対する予測性能が，他の 2 手法と比較して高いことを確認した. 2つの要因がこの結果に影響したと考えられる. 1つ目は, CGP の柔軟性である. CGP は MLR と比較して柔軟な予測モデルを構築する. そして2つ目は, 予測モデルのシンプルさである. この特徴によって, 教師データの過学習が抑制されるため, 未知データに対して, 高い予測精度が得られたと推測できる.

6.5.2 紙の特性へ影響を与えるその他の要因

本研究ではパルプ紙料中の繊維の特徴量を利用して紙の特性を予測した. しかし, 紙の特性は, その他の要因, 例えば, ろ水度, てん料の添加量, 歩留まり向上剤, またはその他のさまざまな製造条件の変動にも当然影響を受ける. 一般的には特徴量数が増えるほど, 予測モデルは複雑になり, 構築が困難となる. しかし, CGP は予測に不要な特徴量を自動的に取捨選択する. この特徴から, CGP は特徴量の次元数が高い場合にも高精度な予測モデルを構築可能と考えられる.

6.5.3 CGP によって選択された特徴量

CGP によって選択された特徴量の一例を表 6.5 に示す. この表では Mean の予測モデルを代表として示す. 各値は, set1 から set5 における各特徴量の使用回数の合計である. 表 6.5 は F01, F02, F08, F22, F26 が, Mean の予測に高い頻度で使用されていることを示している. この結果はこれらの値が Mean の予測に重要であることを示している. 一方, 高い相関関係のある特徴量群を使用したため, 頻度は少ないが, その他の特徴量も広く使用されている.

表 6.6 各繊維特徴量の選択頻度

Table 6.6 Employed frequency of feature amounts of fiber

Frequency	ID
5	F08
4	F01 F02 F22 F26
3	F06 F16 F17 F21 F28 F44
2	F04 F07 F10 F11 F13 F15 F23 F27 F30 F31 F32 F37 F38 F39 F42
1	F05 F12 F14 F20 F24 F25 F29 F34 F35 F41 F43 F46
0	F03 F09 F18 F19 F33 F36 F40 F45 F47

6.5.4 CGP によって構築された予測モデルについて

CGP によって構築されたモデルにおいて、製紙に関する研究者および製造者と意見交換を行ったところ、以下のような意見が得られた。

- NN の構造より特徴量間の関係が理解しやすい。
- 多種類の特徴量を利用した予測性能の向上において期待ができる。
- 既知の概念以外の情報の気づきにおいて有効である。
- MLR のような統計的に確立された手法と比べると信頼性が低い。
- MLR, NN と比べて予測性能は向上しているが、その向上幅が小さいこのため、現時点の性能では、実用面において大きな改善とは言えない。

意見交換の内容を整理すると、CGP によって構築されたモデルは、予測過程の理解のしやすさにおいて一定の効果はあるが、予測性能および予測結果の信頼性において、さらなる改善が必要であるとされた。

6.6 まとめ

本章では、パルプ紙料から得られる情報を基に、紙の品質を予測する手法を検討した。実験では、CGP を用いた提案手法に加え MLR, NN を用い、パルプ紙料中の繊維から得られる 47 種類の特徴量から、紙の光学特性の予測を行った。サンプルには 9 種類のパルプ紙料および同パルプ紙料から作製した手すきシートを用い、各手法の予測精度および予測モデルの構造について評価を行った。その結果、CGP によって得られた予測モデルの汎化性能は、MLR および NN と比較して高かった。また、CGP の次元削減能力は MLR と比較して高かった。以上の結果から、紙の特性の予測において CGP を用いた提案手法の有効性を見出した。

第7章 総括

7.1 本論文で得られた成果

本論文では、貴重印刷物用紙の製造において、紙の特性の管理方法の高度化を図ることを目的に、機械学習を用い、高い分類性能および予測性能と人に理解し易い構造を持つモデルを構築する手法を検討した。各章で得られた成果は以下のとおりである。

① 複数の画像特徴量と FS-SVM による紙の分類モデルの自動構築

製造条件の異なる紙の分類において、画像から算出した複数の画像特徴量の最適化と FS-SVM によって、自動的に紙の分類器を構成する方法を提案した。実験結果では、メーカーの異なる紙およびワイヤーの異なる手すきシートの分類に加えて、その他の様々な条件が異なる手すきシートに対しても高精度な分類器を自動的に構成することができ、また多数の特徴量の中から分類に必要な特徴量を明確にすることが可能であった。以上から提案手法の有効性を示した。

② 複数の画像特徴量と EDEN による紙の分類モデルの自動構築

分類過程の可読性と高い分類精度を両立した紙の分類モデルを構築することを目的に EDEN を用いた紙の分類モデルの自動構築方法を提案した。EDEN の構造においては、単一のデータ群における特徴量間の相互作用の表現を実現するためにノードを拡張し、複数の特徴量を合成する処理を実装した。実験においては、EDEN を用いて構成した分類モデルは多様な紙の分類において高い分類性能を示し、また、分類過程の可読性においても簡潔なモデルを構築したことから提案手法の有効性を示した。

③ パルプ紙料中の複数の繊維特徴量と CGP による紙の予測モデルの自動構築

紙の特性の予測において、高い予測精度と予測過程の可読性の高いモデルを構築することを目的に、CGP を用いた紙の特性の予測モデルの自動構築方法を提案した。CGP の構造においては、整数値のみでは表現の難しい処理を実現するためにノードにおける処理を拡張し、各ノードが個別に実数値を用いた処理を表現可能な構造とした。実験では、パルプ紙料中の繊維から得られる複数の特徴量から紙の光学特性を従来手法より高精度に予測し、また、簡潔なモデルを構築したことから提案手法の有効性を示した。

以上の成果によって、紙の特性管理の高度化に求められる分類・予測精度の向上およびそのモデルの可読性の向上において、機械学習を用いた手法の有効性を示した。

7.2 今後の課題

本研究をさらに発展させるための課題として、以下の点が挙げられる。

① より現実問題に即した条件での検証

本論文において実施した実験では、研究室レベルで作成した手すきシートやメーカーの異なる紙を使用し、紙の特性として大きく異なるものをサンプルとして主に取り扱った。また、紙料を構成する要素のうち、繊維の特徴量のみに着目した。今後の発展として、本論文における基礎検証の成果を基に実際の製造工程におけるより詳細な紙の特性および紙料の特性の変動に対する検証が必要である。

② 製紙工程全体を考慮したシステムの構築

本論文では、紙の光学特性とパルプ紙料における繊維特徴量を検証の対象として主に取り扱ったが、第1章および第2章で示したように紙の特性は多様で、紙の製造工程は非常に複雑である。本論文での基礎検証の成果を生かし、さらなる発展として図7.1に示すような原料から各工程におけるすべての要因を考慮した管理システムを構築することが必要とされる。このように複雑で大規模な問題になる程、機械学習を用いた手法の有効性を活用することが重要である。

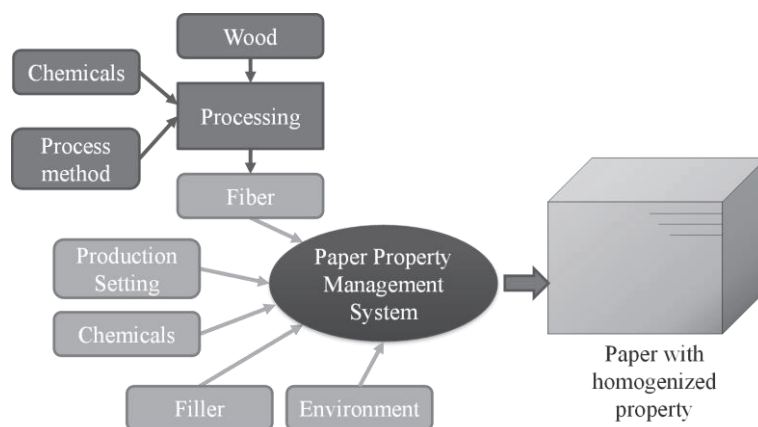


図 7.1 原料から製造工程までのすべての要素を考慮した予測システム

Fig. 7.1 Prediction system with all factors (from Material to all manufacturing processes)

製造現場における実問題への適用およびより実用的なシステムの構築において、以上の課題に取り組む必要がある。

謝辞

平成 23 年から 4 年間の長きに当たり，親身なご指導と適切なお助言を賜りました長尾智晴先生に深く感謝を申し上げます。

本論文をまとめるにあたり，貴重なご指導，ご助言を頂きました，横浜国立大学環境情報研究院，田村直良先生，森辰則先生，岡嶋克典先生，富井尚志先生に深く感謝を申し上げます。

また，本研究は，長尾研究室の皆様，千葉大学大学院，矢田紀子先生，国立印刷局研究所の同僚の皆様からのご教授とご協力によって完成することができました。厚く謝意を表します。

そして，仕事と学業の両立において，精神的に支えてくれた妻に心から感謝します。妻の信頼と支えがあったからこそ，本研究を最後までやり遂げることができました。

最後に，研究テーマとの出会い，研究仲間との出会い，そしてご指導を賜りました素晴らしい先生方との出会いに深く感謝して本論文を結びます。

平成 27 年 3 月 石川 卓

参考文献

- [1] 佐伯仁志, “通貨偽造罪の研究,” 金融研究, 2004.
- [2] "CEN Workshop 6 Security Management System for Secure Printing," [Online]. Available: <http://www.cen.eu/cen/Secrors/TechnicalCommitteesWorkshop/6>.
- [3] 平成 26 年版金融情報システム白書, 経産省, 2013.
- [4] 日本銀行, “決済システムレポート,” 2013.
- [5] 警視庁, “平成 20 年度警察白書,” 2008.
- [6] 警視庁, “平成 26 年度警察白書 統計資料,” 2014.
- [7] 財務省印刷局, “連携 IC カード券面の偽造防止技術ハンドブック,” 2002.
- [8] R. L. v. Renesse, "Optical Document Security, Third Edition," 2005.
- [9] 山越学, “人工物メトリクスによる印刷基材の認証,” 博士論文, 横浜国立大学, 2012.
- [10] 国立印刷局, “お札の偽造防止技術～現在発行されているお札～,” [オンライン]. Available: <http://www.npb.go.jp/ja/intro/gizou/genzai.html>.
- [11] 国立印刷局, “お札の製造工程,” [オンライン]. Available: <http://www.npb.go.jp/ja/intro/seizou.html>.
- [12] 日本自動販売機工業会, “自販機普及台数及び年間自販金額,” 2013.
- [13] 電子情報技術産業協会, “端末装置に関する調査報告書,” 2014.
- [14] 縄田和満, “回帰分析,” 著: 統計学入門, 東京大学出版会, 2004, pp. 258-276.
- [15] JIS P 0001:1998, *Paper, board and pulp – Vocabulary*.
- [16] 門屋卓, “印刷・情報記録用紙としての基本特性,” 著: 印刷・情報記録における紙の特性と印刷適性および分析, 評価, 技術情報協会, 1999, pp. 3-34.
- [17] 紙パルプ技術協会出版委員会, “紙の分類と規格,” 著: 紙パルプ技術便覧, 紙パルプ技術協会, 1992, pp. 455-502.
- [18] 磯貝明, “紙の化学分析と機能発現機構,” 著: 印刷・情報記録における紙の特性と印刷適性および分析, 評価, 技術情報協会, 1999, pp. 77-81.
- [19] 東京都立図書館, “酸性紙思量の脱酸性化処置,” [オンライン]. Available: http://www.library.metro.tokyo.jp/about_us/syusyu_hozon/siryu_hozon/tabid/2126/Default.aspx.
- [20] 国立印刷局, “お札の特徴,” [オンライン]. Available: <http://www.npb.go.jp/ja/intro/tokutyou/index.html>.

- [21] Bank of England, "Banknote substrates," [Online]. Available: <http://www.bankofengland.co.uk/banknotes/polymer/Documents/banknotesubstrates.pdf>.
- [22] 紙パルプ技術協会出版委員会, “紙パルプの物性と試験法,” 著: 紙パルプ技術便覧, 紙パルプ技術協会, 1992, pp. 413-453.
- [23] 門屋卓, 角祐一郎, 吉野勇, 新・紙の科学, 中外産業調査会, 1989.
- [24] 江前敏晴, “紙パルプ試験規格委員会の活動と JIS・ISO 規格の最新動向,” 紙パ技協誌, 第 57 巻, 第 2 号, pp. 240-250, 2003.
- [25] 横溝秀なお, 木村実, “紙の物理的試験法,” 著: 紙パルプの試験法, 紙パルプ技術協会, 1995, pp. 103-212.
- [26] JIS P 8111:1998, *Paper, board and pulps – Standard atmosphere for conditioning and testing.*
- [27] JIS P 8124:2011, *Paper and board-Determination of grammage.*
- [28] JIS P 8118:1998, *Paper and board – Determination of thickness and density.*
- [29] JIS P 8133:2013, *Paper, board and pulps-Determination of pH of aqueous extracts-Part 1: Cold extraction.*
- [30] JIS P 8133-2:2013, *Paper, board and pulps-Determination of pH of aqueous extracts-Part 2: Hot extraction.*
- [31] JIS P 8127, *Paper and board-Determination of moisture content of a lot-Oven-drying method.*
- [32] JIS P 8251:2003, *Paper, board and pulps-Determination of residue (ash) on ignition at 525 degrees C.*
- [33] JIS P 8113:2006, *Paper and board-Determination of tensile properties-Part 2: Constant rate of elongation method.*
- [34] JIS P 8115:2001, *Paper and board – Determination of folding endurance – MIT method.*
- [35] 村上浩二, “紙層構造の評価とその進歩,” 第 40 巻, 第 12 号, pp. 1089-1100, 1986.
- [36] 湯原隆, 蓮池牧雄, 村上浩二, “コンピュータ画像処理によるパルプシートの構造解析 (第 1 報) 2 次元的质量変動の定量的評価法について,” 紙パ技協誌, 第 40 巻, 第 89 号, pp. 85-91, 1986.
- [37] 湯原隆, 蓮池牧雄, 村上浩二, “同時生起行列による紙の地合い評価,” 繊維学会誌, 第 42 巻, 第 5 号, pp. 279-284, 1986.
- [38] 湯原隆, 蓮池牧雄, 村上浩二, “コンピュータ画像処理によるパルプシートの構造解析 (第 2 報) 紙の地合い評価,” 紙パ技協誌, 第 40 巻, 第 10 号, pp. 949-958, 1986.
- [39] R. J. Allan. and M. C. Shingleton, "The Development of an Ultrasonic Measurement for the Prediction of Paper Strength Properties - Part2," *Appita Conference and Exhibition*, pp. 363-373, 2008.

- [40] 滝. 政幸, 小. 幸夫, 井上信一, “紙の内部散乱光特性の評価,” *紙パ技協誌*, 第 卷 63, 第 1, pp. 54-56, 2010.
- [41] R. Danby and H. Zhou, "Numerical evaluation of the printability of paper surfaces," *Pulp & Paper Canada*, vol. 105, no. 9, pp. 38-42, 2004.
- [42] H. Hagglund, O. Norberg and P. Edstrom, "Prediction of optical variations in paper from high resolution measurements of paper properties," *Nordic Pulp & Paper Research Journal*, vol. 28, no. 4, pp. 596-601, 2013.
- [43] D. K. Casper, M. Penetti and E. H. Kim, "Image analytical monitoring of paper quality - a feasibility study," *TAPPI Journal*, vol. 5, no. 11, pp. 18-24, 2006.
- [44] A. Dilley, "Formation Measurement to Aid Quality Control," *Conference Proceeding PITA Annual Conference*, pp. 43-47, 1996.
- [45] 門屋卓, 白田誠人, 清水弘, “光学的方法によるパルプ繊維のフィブリル化の評価,” *紙パ技協誌*, 第 卷 35, 第 7, pp. 609-615, 1981.
- [46] M. Williamson, "Kajaani Fiber Analysis Products: Laboratory quality precision with online analyzer reliability," *results automation customer magazine*, no. 4, pp. 8-11, 2006.
- [47] 尾崎史彦, “紙設計の基礎としての繊維特性計測,” *紙パルプ技術タイムス*, 第 12, pp. 45-50, 2006.
- [48] JIS P 8121-1:2012, *Pulps-Determination of drainability-Part 1: Schopper-Riegler method.*
- [49] JIS P 8121-2:2012, *Pulps-Determination of drainability-Part2:"Canadian Standad" freeness method.*
- [50] TAPPI T 232:2013, *Fiber length of pulp by profection.*
- [51] JIS P 8226, *Pulps-Determination of fibre length by automated optical analysis-Part 1: Polarized light method.*
- [52] JIS P 8226-2:2011, *Pulps-Determination of fibre length by automated optical analysis-Part 2:Unpolarized light method.*
- [53] ISO 16065-1:2014, *Pulps -- Determination of fibre length by automated optical analysis -- Part 1: Polarized light method.*
- [54] J. Ruonala, I. Joensuu and T. Niskanen, "Analysis Control of Pulp Quality and Prediction of Paper Quality Using New Imaging Technology," *紙パルプ技術タイムス*, 2012.
- [55] N. Jan, C. E. Johan and L. Torbjoern, "Joint Estimation of Fibers and Fines Concentration in Paper Pulp Suspensions Using a Combined Optical and Acoustic Technique," *IEEE International Ultrasonics Symposium*, vol. 2009, no. 2, pp. 1191-1194, 2009.
- [56] P. Iiro and A. Ville, "New model for predicting tensile strength and density of eucalyptus

- handsheets based on an activation parameter calculated from fiber distribution characteristics," *Holzforschung*, vol. 64, no. 2, pp. 201-209, 2010.
- [57] J. L. Shi, S. Y. Zhang and B. Riedl, "Multivariate modeling of MDF panel properties in relation to wood fiber characteristics," *Holzforschung*, vol. 60, no. 3, pp. 285-293, 2006.
- [58] L. Randy, F. Brenton and C. Devin, "Paper quality real time optimisation," *Appita Journal*, vol. 58, no. 5, pp. 346-348, 2005.
- [59] K. R. Braaten and D. Molteberg, "A mathematical method for determining fiber wall thickness and fiber width," *TAPPI Journal*, vol. 3, no. 2, pp. 9-12, 2004.
- [60] H. Juha, S. Pooya and K. Pasi, "Automatic image-based detection of paper fiber ends," *Proceedings of SPIE*, vol. 8009, pp. 80092N.1-80092N.5, 2011.
- [61] JIS P 8222:1998, *Pulps – Preparation of laboratory sheets for physical testing*.
- [62] ISO 5269-1:2005, *Pulps -- Preparation of laboratory sheets for physical testing -- Part 1: Conventional sheet-former method*.
- [63] T. Back, F. Hoffmeister and H.-P. Schwefel., "A survey of evolution strategies," *Proceedings of the Fourth International Conference on Genetic Algorithms*, pp. 2-9, 1991.
- [64] 長尾智晴, “進化的計算法,” 著: 最適化アルゴリズム, 昭晃堂, 2000, pp. 136-150.
- [65] J. H. Holland, "Adaptation in natural and artificial systems," *Ann Arbor: University of Michigan*, 1975.
- [66] D. E. Goldberg, "Genetic algorithm in search," *Optimization and Machine Learning*, 1989.
- [67] 安居院猛, 長尾智晴., ジェネティックアルゴリズム, 昭晃堂, 1993.
- [68] J. R. Koza, "Genetic programming: on the programming of computers by means of natural," *MIT Press*, 1992.
- [69] 伊庭齊志, “遺伝的プログラミング,” 1996.
- [70] 白川真一, 矢田紀子, 長尾智晴, “遺伝的プログラミングによる実数値交叉の性能差を強調する探索空間の生成,” *進化計算学会論文誌*, 第 1 巻, 第 1, pp. 54-64, 2010.
- [71] L. J. Echelman and J. D. Schaffer, "Real Coded Genetic Algorithms and Interval-Schemata," in *Foundations of Genetic Algorithms 2*, Morgan Kaufmann Publishers, 1993, pp. 187-202.
- [72] 小野功, 佐藤浩, 小林重信, “単峰性正規分布交叉 UNDX を用いた実数値 GA による関数最適化,” *人工知能学会誌*, 第 14 巻, 第 6, pp. 1146-1155, 1999.
- [73] 樋口隆英, 筒井茂義, 山村雅幸, “実数値 GA におけるシンプレクス交叉の提案,” *人工知能学会論文誌*, 第 16 巻, 第 1, pp. 147-155, 2001.
- [74] C. M. Bishop, “ニューラルネットワーク,” 著: パターン認識と機械学習 上, シュプリンガー・ジャパン, 2007, pp. 225-287.

- [75] D. E. Mumelhart, G. E. Hinton and R. J. Williams, "Learning Internal Representation by Error Propagation," *Parallel distributed processing: explorations in the microstructure of cognition*, MIT Press, vol. 1, pp. 318-362, 1986.
- [76] 長尾智晴, “誤差逆伝搬法,” 著: 最適化アルゴリズム, 昭晃堂, 2000, pp. 104-119.
- [77] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," in *Proceedings of the National Academy of Sciences U. S. A.*, vol. 79, pp. 2554-2558, 1982.
- [78] J. J. Hopfield and T. W. Tank, "Neural Computation of Decisions in Optimization Problems," *Biological Cybernetics*, vol. 52, pp. 141-152, 1985.
- [79] C. M. Bishop, “疎な解を持つカーネルマシン,” 著: パターン認識と機械学習 下, シュプリンガー・ジャパン, 2008, pp. 35-69.
- [80] N. Cristianini, J. Shawe-Taylor, 訳: 大北剛, サポートベクターマシン入門, 共立出版, 2005.
- [81] A. Nicolas, C. Hyun-Woo, B. S. Hyun, J. Myong-Kee and Y. T. M., "Prediction of internal bond strength in a medium density fiberboard process using multivariate statistical methods and variable selection," *Wood Science Technology*, vol. 42, no. 7, pp. 521-534, 2008.
- [82] O. Philippa, M. Sayyad and M. A. J., "Exploring data for relationships between wood, fibre and paper properties," *Appita Journal*, vol. 52, no. 5, pp. 358-362, 382, 1999.
- [83] R. Alejandro, S. Rafael, F. Ana and R. Ana, "Simulation of Hesperaloe funifera diethanolamine pulping by polynomial and neural fuzzy models," *Chemical Engineering Research and Design*, vol. 89, no. 6, pp. 648-656, 2011.
- [84] T. Miyanishi and H. Shimada, "Using neural networks to diagnose web breaks on a newsprint paper machine," *TAPPI Journal*, vol. 81, no. 9, pp. 163-170, 1998.
- [85] A. Kumar and V. C. Hand, "Using Genetic Algorithms and Neural Networks To Predict And Optimize Coated Board Brightness," *TAPPI Papermarks Conference Proceedings*, vol. 1, pp. 161-170, 1999.
- [86] R. B. Chinnam, "Support vector machines for recognizing shifts in correlated and other manufacturing processes," *International Journal of Production Research*, vol. 40, no. 17, pp. 4449-4466, 2002.
- [87] 宮田瞳, 篠崎真, “フーリエ変換と相互相関法による紙の異動識別法,” *紙パ技協誌*, 第 54 巻, 第 3 号, pp. 396-401, 2000.
- [88] 篠崎真, 宮田瞳, 中山智仁, 江前敏晴, “フーリエ変換と相互相関法による紙の異動識別法 (第 2 報) 市販 PPC 用紙への適用,” *紙パ技協誌*, 第 55 巻, 第 4 号, pp. 514-521, 2001.

- [89] 宮田瞳, 下山昌彦, 篠崎真, “レーザ変位計と画像処理を用いた画用紙の異動識別法,” *紙パ技協誌*, 第 57 巻, 第 5 号, pp. 724-731, 2003.
- [90] K. I. Kim, K. Jung, S. H. Park and H. J. Kim, "Support Vector Machines for Texture Classification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1542-1550, 2002.
- [91] 西田健次, 栗田多喜夫, “カーネル学習法とその画像認識への応用,” *情報処理学会論文誌*, 第 46 巻, 第 SIG15 (CVIM12), pp. 1-10, 2005.
- [92] 余部治昭, 長尾智晴, “特徴量選択型 SVM を用いた欠陥画像分類,” *情報科学技術フォーラム*, 第 巻 FIT2006, pp. 79-80, 2006.
- [93] デジタル画像処理編集委員会, 著: *デジタル画像処理*, 画像情報教育振興協会, 2004, pp. 175-176.
- [94] M. Petrou and P. G. Sevilla, in *Image Processing Dealing With Texture*, JOHN WILEY & SONS INC, 2006, pp. 275-295.
- [95] J. R. Quinlan, "Improved Use of Continuous Attributes in C4.5," *Journal of Artificial Intelligence Research*, vol. 4, pp. 77-90, 1996.
- [96] 中山史朗, 穂積知佐, 矢田紀子, 長尾智晴, “進化的条件判断ネットワークによる画像分類,” *知能システムシンポジウム資料*, 第 38 巻, pp. 277-282, 2011.
- [97] 佐藤博, 小野功, 小林重信, “遺伝的アルゴリズムにおける世代交代モデルの提案と評価,” *人工知能学会誌*, 第 12 巻, 第 5 号, pp. 734-743, 1996.
- [98] 小林重信, “実数値 GA のフロンティア,” *人工知能学会誌*, 第 24 巻, 第 1 号, pp. 147-162, 2009.
- [99] A. Goran, "Fundamentals of Pulp Fiber Quality and Paper Properties," *TAPPI Proceedings Pulping Conference*, vol. 1999, no. 1, pp. 29-39, 1999.
- [100] J. F. Miller and P. Thomson, "Cartesian Genetic Programming," *Genetic Programming: Lecture Notes in Computer Science*, vol. 1802, pp. 121-132, 2000.
- [101] K. Volk, J. F. Miller. and S. L. Smith, "Multiple Network CGP for the Classification of Mammograms," *Applications of Evolutionary Computing: Lecture Notes in Computer Science*, vol. 5484, pp. 405-413, 2009.
- [102] 宮内崇 and 長尾智晴, "Cartesian Genetic Programming を用いた医用画像分類のための特徴量自動構築," *知能システムシンポジウム資料*, vol. 40 巻, pp. 175-178, 2013.
- [103] H. Akaike, "A new look at statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716-723, 1974.

研究業績

論文（国内）

[Pub. 1] 石川卓, 矢田紀子, 長尾智晴: 画像特徴量の組み合わせ最適化による紙の自動分類, 紙パルプ技術協会誌, Vol. 65, No. 6, pp. 49-58 (2011).

論文（海外）

[Pub. 2] Taku Ishikawa, Takashi Okuda, Tomoharu Nagao: Creation of Paper Property Prediction Models Using Cartesian Genetic Programming, *Appita Journal* (2015). *In Press*

翻訳論文

[Pub. 3] Taku Ishikawa, Noriko Yata, Tomoharu Nagao: Automatic Classification of Paper Using Combinational Optimization of Image Features, *Japan TAPPI Journal*, Vol. 65, No. 6, pp. 59-68 (2011). ([Pub.1]の翻訳)

学会発表

[Pub. 4] 石川卓, 長尾智晴: 進化型条件判断ネットワークによる紙の分類, 第 78 回紙パルプ研究発表会 (2011).