

パノラミック検索モデルに基づく
インタラクティブなマルチメディア検索
および解析手法とその応用



東京大学
THE UNIVERSITY OF TOKYO

情報理工学系研究科 電子情報学専攻

長田 茂美

Committee:

相澤 清晴 (主査)

安達 淳

石塚 満

喜連川 優

豊田 正史

Supervisor:

池内 克史

論文要旨

インターネットの爆発的な拡がり，ブロードバンド化の進展により，社会活動のあらゆる局面で大量かつ多種多様な情報が創出，蓄積されている．これに伴い，大量かつ多種多様な情報の中から必要な情報を効率的に検索したり，それらの情報を解析したり，あるいは，その中に埋もれている有用な知識を発見するといった情報を活用するための基盤となる次世代検索・解析技術が注目を集めている．近年，特に，計算機処理能力の飛躍的な向上も相俟って，画像，音声，テキストなど多様なメディアが混在するマルチメディア・コンテンツの創出，蓄積も爆発的に増加しており，マルチメディア・コンテンツの検索・解析技術への期待が益々高まっている．

一般に，マルチメディア・コンテンツは，画像，音声，テキストなどのメディアが相互に関連性を持って混在する複合メディアの形態をとる．このようなマルチメディア・コンテンツの検索には，従来から，マルチメディア・コンテンツを構成する各メディアに人手でメタデータを付与し，あるいは，コンテンツ内のテキストをメタデータとして取り扱い，そのメタデータを対象としたキーワード検索により，所望のコンテンツを検索する手法が採られてきた．近年では，マルチメディア・コンテンツを構成する各メディアの内容に関する特徴を自動的に抽出し，その特徴に基づいて検索する手法（Content-Based Multimedia Information Retrieval），例えば，画像の色や形状などの画像特徴を用いる“内容に基づく画像検索手法（Content-Based Image Retrieval，CBIR）”の適用も進められている．

しかしながら，これらの検索手法は，ある特定の観点に基づいて付与あるいは抽出したメタデータを利用し，マルチメディア・コンテンツを検索するものであり，メタデータが画像の色や形状などの低次画像特徴である場合には，システムが抽出できる低次画像特徴と人間の検索意図である高次の意味概念との間のいわゆる Semantic Gap の問題が生じる．

本論文では，この Semantic Gap の解消を目指し，人間を積極的に検索および解析ループの中に取り込むことにより，人間とシステムとのインタラクションを通じて，互いの優れた能力を提供し，支援し合えるインタラクティブなマルチメディア・コンテンツの検索および解析手法を提案する．

具体的には，主に画像とテキストから構成されるマルチメディア・コンテンツを対象として，まず初めに，人間の視覚認識能力を最大限活用するためにシステムが何を支援すべきかを検討し，1) クロスメディア検索，2) 検索結果および検索過程の仮想三次元空間による視覚化，を特徴とするパノラミック検索手法を提案する．提案手法は，マルチメディ

ア・コンテンツを構成する各メディアの特性および各メディア間の関連性を有効に活用し、コンテンツが画像とテキストで構成される場合には、画像による視覚的な検索とテキストによる意味的な検索とを兼ね備えた情報検索手段を提供する。また、検索結果としてのコンテンツ集合やインタラクティブな検索過程を仮想三次元空間上で人間に分かりやすく視覚化することによって、人間の視覚認識能力を活用した効率的な検索が実現できる。

次に、パノラミック検索手法の枠組みを、大量のマルチメディア・コンテンツからの知識発見を支援する解析手法の枠組みへと拡張し、画像群からの知識発見を支援するビジュアル解析手法として、パノラミック解析手法を提案する。画像群からの知識発見では、無数に定義できる視覚的特徴がある中で、有用な知識に關与する視覚的特徴を発見することが本質的な課題である。提案手法は、人間の視覚認識能力を有効に活用し、人間による知識発見を支援する立場を採ることにより、インタラクティブに有用な知識に關与する視覚的特徴さらには有用な知識を見出すことが可能となり、応用の幅を拡げることができる。

さらに、人間とシステムとのインタラクションの一つとして、適合フィードバックを導入し、能動学習に基づく適合フィードバックの新たな手法を提案する。提案手法は、適合フィードバックの反復が少ない段階においても、人間の検索意図を反映した検索を効率的に実現できる。

本論文の主要な貢献は、マルチメディア・コンテンツの検索および解析に人間とシステムとのインタラクションを活用する三つの手法を提案したことである。具体的には、1) インタラクティブなマルチメディア・コンテンツの検索手法であるパノラミック検索手法、2) パノラミック検索手法を拡張した画像群からの知識発見を支援するパノラミック解析手法、3) 人間の検索意図に効率的に適合する能動学習に基づく適合フィードバック手法、を提案し、評価実験により、提案手法の有効性を確認した。また、これらの提案手法を適用した実システムを開発し、設計・製造、企業内情報共有の分野などの実応用においても提案手法の有効性を確認できた。

Abstract

With the explosive expansion of the Internet and the broadband capability, a wide variety of information has been created and stored by all phases of social activities. Therefore, being able to efficiently retrieve information from the large pool of data, as well as being able to discover useful knowledge buried in the data, have become the essential focus of the next generation retrieval and analysis technology. In recent years, the rapid improvement in the computer processing capabilities has led to the dramatic growth in the creation and storage of multimedia contents, which are composed of various media like image, sound and text, etc. The retrieval and analysis technology of multimedia contents has thus become an active and promising field of study.

In general, multimedia contents are composed of two or more media such as the image, sound and text, etc., which are associated with each other. The conventional methods in multimedia content retrieval have been either adding the metadata manually to each media, or using the text composed in the multimedia content as the metadata and retrieves the content via keyword search. In recent years, Content-Based Multimedia Information Retrieval, which extracts the features of media content, has become a fast growing field of study. The technique of Content-Based Image Retrieval (CBIR), for example, uses the visual features of image, such as the color and shape, to retrieve the data.

However, among these retrieval techniques, when the system retrieves multimedia contents by using the metadata with low-level image features like color or shape, the problem of so-called Semantic Gap occurs between the low-level image feature and the high-level semantic concepts of user's retrieval intention.

This research aims to resolve the Semantic Gap problem by proposing the interactive retrieval and analysis methods of multimedia contents. By actively involving the user in the process of retrieval or analysis, through interaction between the user and the system, both are able to provide and support each other at great efficiency.

This thesis targets mainly multimedia contents composed of image and text. It firstly examines how the system should support to maximize the human's visual recognition ability, which the panoramic retrieval method is presented featuring 1) cross-media retrieval and 2) visualization of the retrieval result and retrieval process in a virtual 3D space. The method effectively utilizes the feature of each media

contained in the multimedia content and also the association between each media; which in the case of multimedia contents composed of image and text, it offers a combination of visual retrieval via image and semantic retrieval via text. Moreover, by visually presenting the retrieval result of contents collection and the interactive search process in a virtual 3D space, the retrieval is achieved efficiently via the human's visual recognition ability.

Next, by enhancing the framework of the panoramic retrieval method with analysis ability that supports the knowledge discovery from a large amount of multimedia contents, the panoramic analysis method is introduced for analyzing the data visually. Within a group of images, where a countless number of image features can be defined, the challenge lies in being able to discover the image feature that provides useful knowledge. Through utilizing the human's visual recognition ability effectively, as well as supporting the human activity in analysis, this method is able to discover the image feature with useful knowledge via user interaction, and can be expanded in various applications.

In addition, as one of the interactions between the user and the retrieval system, a new method adopting relevance feedback based on active learning is introduced. The proposed method efficiently achieves the result that reflects user's retrieval intention even at the stage with only a few repetitions of relevance feedback.

The main contribution of this thesis is to propose three methods using the interaction between the user and the system to retrieve and to analyze the multimedia contents. The three methods are: 1) The panoramic retrieval method for interactive multimedia contents retrieval. 2) The panoramic analysis method that supports the knowledge discovery from images as an enhancement to the panoramic retrieval technique. 3) Active feedback that efficiently reflects the user's retrieval intention. Experiment results have proved the effectiveness of the three methods. In addition, actual applications of the proposed methods in various fields, such as design & manufacturing, and data sharing within the corporate environment, have further confirmed the effectiveness of the methods.

目次

論文要旨	i
Abstract	iii
目次	vi
図目次	ix
表目次	xi
第 1 章 序論	1
1.1 本論文の背景	1
1.2 本論文の目的	3
1.3 本論文の構成	3
第 2 章 パノラミック検索モデルに基づくインタラクティブなマルチメディア検索手法	5
2.1 はじめに	5
2.2 パノラミック検索モデル	5
2.3 パノラミック検索手法	8
2.4 Web 検索システム	9
2.5 映像検索システム	17
2.6 その他の応用	20
2.7 関連研究	30

2.8	実験	31
2.9	まとめ	42
第 3 章	検索手法から解析手法への拡張	43
3.1	はじめに	43
3.2	関連研究	43
3.3	パノラミック解析手法	44
3.4	実験	47
3.5	まとめ	56
第 4 章	適合フィードバックの導入	57
4.1	はじめに	57
4.2	関連研究	58
4.3	Representative 基準を持つ SVM の能動学習	59
4.4	提案手法の枠組み	66
4.5	実験	68
4.6	まとめ	81
第 5 章	結論	83
5.1	本論文の主たる成果	83
5.2	今後の課題と展望	84
謝辞		85
発表文献		87
参考文献		97

2.1	仮想三次元空間と検索過程の視覚化	6
2.2	パノラミック検索手法の構成	8
2.3	Web 検索システム	9
2.4	画像と関連テキスト	10
2.5	画像の色特徴による一覧表示平面の表示例	12
2.6	テキスト特徴による一覧表示平面の表示例	12
2.7	選別による絞込み検索結果の表示例	13
2.8	排除による絞込み検索結果の表示例	14
2.9	画像の形状特徴による類似コンテンツリストの表示例	15
2.10	LOD 表示	16
2.11	映像検索システム	17
2.12	シーン代表画像の色特徴による表示例	19
2.13	シーン代表画像のテキスト特徴による表示例	19
2.14	デジカメ画像検索システム	21
2.15	デジタル文書検索システム	22
2.16	図面検索システム	23
2.17	3D-CAD モデル検索システム	24
2.18	3D-CAD モデルの Shape Distribution 特徴による表示例	25
2.19	ハイブリッド・レンダリング	25
2.20	ハイブリッド・レンダリングによる 3D-CAD モデルの表示例	26

2.21	ポーズ・アラインメントの仮定	27
2.22	大局的ポーズ・アラインメント	28
2.23	局所的ポーズ・アラインメント	29
2.24	局所的ポーズ・アラインメントの類似度計算	29
2.25	デジタル文書検索実験 (Task2)	32
2.26	デジタル文書検索実験 (Task3)	33
2.27	デジタル文書検索実験の結果	34
2.28	映像検索実験の標的画像	36
2.29	HSI 色ヒストグラム特徴に基づく配置と標的画像 B	36
2.30	映像検索実験の結果 (配置方法による比較)	37
2.31	映像検索実験の結果 (3 倍速再生との比較)	37
2.32	画像検索実験の標的画像	39
2.33	画像検索実験 (Task1)	40
2.34	画像検索実験の結果	41
3.1	パノラミック解析手法の概要	45
3.2	テキスト特徴による陶磁器画像の配置	48
3.3	重要単語のリスト表示	49
3.4	“ 染付 ” による絞込み検索結果	49
3.5	関連データ “ VF8 ” による配置	50
3.6	“ MG1 ” 層の曲率と “ VF8 ” との関係	51
3.7	TMR ヘッドとその CAE 画像	52
3.8	CAE 画像と電圧波形の “ 非対称度 ”	53
3.9	関連データ “ 非対称度 ” による配置	53
3.10	CAE 画像の色特徴および絞込み検索による配置	55
4.1	Representative 基準	61
4.2	提案手法の枠組み	67
4.3	機械部品画像の例	69

4.4	機械部品画像（反復回数=0）	71
4.5	機械部品画像（反復回数=0，適合／非適合の評価）	72
4.6	機械部品画像（反復回数=1）	73
4.7	ラベルなし画像の活用／非活用（機械部品画像）	75
4.8	ラベルなし画像の活用／非活用（Corel 画像）	76
4.9	λ の影響（ $L^* = 6$ ）	77
4.10	Angle-Diversity 手法との比較（機械部品画像）	79
4.11	Angle-Diversity 手法との比較（Corel 画像）	80

表目次

2.1	パノラミック検索手法による検索対象の表現図式	20
4.1	Normalized Cuts 法による動的クラスタリングの実行時間	77

第 1 章

序論

1.1 本論文の背景

インターネットの爆発的な拡がり，ブロードバンド化の進展により，社会活動のあらゆる局面で大量かつ多種多様な情報が創出，蓄積されている．これに伴い，大量かつ多種多様な情報の中から必要な情報を効率的に探索したり，解析したり，あるいは，その中に埋もれている有用な知識を発見するといった情報を活用するための基盤となる次世代検索・解析技術の重要性も益々高まっている．

近年，特に，計算機処理能力の飛躍的な向上も相俟って，画像，音声，テキストなど多様なメディアが混在するマルチメディア・コンテンツの創出，蓄積も爆発的に増加しており，その活用基盤としてのマルチメディア・コンテンツの検索・解析技術の研究開発が活発化してきている [1]．

一般に，マルチメディア・コンテンツは，画像，音声，テキストなどのメディアが相互に関連性を持って混在する複合メディアの形態をとる．このようなマルチメディア・コンテンツの検索には，従来から，マルチメディア・コンテンツを構成する各メディアに人手でメタデータを付与し，あるいは，コンテンツ内のテキストをメタデータとして取り扱い，そのメタデータを対象としたキーワード検索により，所望のコンテンツを検索する手法が採られてきた．近年では，マルチメディア・コンテンツを構成する各メディアの内容に関する特徴を自動的に抽出し，その特徴に基づいて検索する手法（Content-Based Multimedia Information Retrieval）[2]，例えば，画像の色や形状などの画像特徴を用いる“内容に基づく画像検索手法（Content-Based Image Retrieval，CBIR）”の適用も進められている [3]，[4]．

しかしながら，これらの検索手法は，ある特定の観点に基づいて付与あるいは抽出した

メタデータを利用してマルチメディア・コンテンツを検索するものであり，ユーザがメタデータの性質を理解した上で，所望のコンテンツを特定する検索クエリを的確に与えなければ，高精度な検索は期待できない．例えば，メタデータが画像の色や形状などの低次画像特徴である場合には，システムが抽出できる低次画像特徴と人間の検索意図である高次の意味概念との間のいわゆる Semantic Gap が生じる．

この Semantic Gap の橋渡しを実現するための一つのアプローチとして，マルチメディア・コンテンツ集合の視覚化の研究が進められており，例えば，検索結果としての文書集合 [5]，画像集合 [6]，[7]，[8] などを計算機内の仮想二次元あるいは三次元空間上で視覚化する手法が提案されている．しかし，これらの手法では，大量のコンテンツ集合をユーザが把握しやすいように視覚化することに主眼が置かれており，視覚化されたコンテンツ集合の中から所望のコンテンツをインタラクティブに絞り込むための検索やその視覚化は十分には扱われていない．

また，Semantic Gap を解消するための別のアプローチとして，検索結果に対するユーザからの適合 / 非適合の評価のフィードバックに基づいて，システムの検索性能をユーザとインタラクティブに改善する適合フィードバック (Relevance Feedback) の研究がある．特に，人間の負荷を軽減し，実応用への適用を目指すという観点から，サポートベクターマシン (Support Vector Machine, SVM) の能動学習 (Active Learning) に基づく適合フィードバック手法が提案されており，さらなる性能の向上，実用性の向上に向けた取り組みが盛んに行われている [9]，[10]，[11]．

一方，マルチメディア・コンテンツから知識を発見するための解析技術としては，人間の脳の f-MRI 画像からの知識発見 [12]，気象画像からの知識発見 [13]，[14]，[15]，医療画像からの知識発見 [16]，航空写真からの知識発見 [17] など，いくつかの分野で，画像群からの知識発見手法が提案されている．しかし，いずれの手法も，発見すべき知識に関与する視覚的特徴が既知であることを前提とし，その視覚的特徴を画像特徴として抽出することにより，データマイニングの問題に帰着させている．現実には，どのような視覚的特徴が発見すべき知識に関与するかが事前に分かっていない問題も数多く存在しているが，このような問題に対処することはできない．また，発見すべき知識に関与する視覚的特徴が既知であったとしても，必ずしもシステムがその視覚的特徴を画像特徴として自動的に抽出できるとは限らない．特に，視覚的特徴がより高次の意味概念レベルであれば，その視覚的特徴を画像特徴として自動的に抽出することは，一般に困難である．ここでも，画像をはじめとするマルチメディア・コンテンツを解釈する上での人間とシステムとの間に存在する Semantic Gap が大きな障壁となっている．

1.2 本論文の目的

本論文では，人間とシステムとの間に存在する Semantic Gap の解消に向けて，人間を積極的に検索および解析ループの中に取り込んだ，インタラクティブなマルチメディア・コンテンツの検索および解析手法を提案するとともに，提案手法を実装した効率的かつ効果的なマルチメディア検索および解析システムを実現することを目的とする．

具体的には，以下の課題に取り組む．

- 人間とのインタラクションを通じて，互いの優れた能力を提供，支援し合えるインタラクティブなマルチメディア・コンテンツの検索手法および解析手法の提案
- 人間とシステムとのインタラクションの一つとして，人間がシステムにフィードバックを与えることにより，システムの自らの学習に基づく能力向上を支援するための新たな適合フィードバック手法の提案

また，これらの手法を適用した実用的なマルチメディア検索および解析システムを構築し，応用面でも貢献する．

1.3 本論文の構成

本論文の構成は以下の通りである．

第 1 章 序論

第 2 章 パノラミック検索モデルに基づくインタラクティブなマルチメディア検索手法

第 3 章 検索手法から解析手法への拡張

第 4 章 適合フィードバックの導入

第 5 章 結論

第 2 章では，マルチメディア・コンテンツのインタラクティブな検索モデル “パノラミック検索モデル” を提案する．パノラミック検索モデルは，マルチメディア・コンテンツを構成する各メディアの特性と各メディア間の関連性を有効に活用し，コンテンツが画像とテキストで構成される場合には，画像による視覚的な検索とテキストによる意味的な検索を兼ね備えた情報検索手段を提供する．この検索モデルを具現化したインタラクティブなマルチメディア検索手法 “パノラミック検索手法” を提案し，Web 検索，映像検索などに適用することにより，提案モデルおよび手法の有効性を論じる．

第 3 章では、第 2 章で提案したパノラミック検索手法の枠組みを、大量のマルチメディア・コンテンツからの知識発見を支援する解析手法の枠組みへと拡張し、画像群からの知識発見を支援するビジュアル解析手法として、“パノラミック解析手法”を提案する。また、パノラミック解析手法を実装したパノラミック解析システムを構築し、教育分野および設計・製造分野における画像群からの知識発見支援に適用することにより、パノラミック解析手法の有効性を論じる。

第 4 章では、SVM の能動学習に基づく適合フィードバック手法を提案し、第 2 章で提案したパノラミック検索手法に導入する。SVM の能動学習に基づく適合フィードバック手法の実現においては、ユーザに適合 / 非適合の評価を求める事例を決定する評価基準が、その手法の性能を左右する重要な鍵となる。本章では、その事例を決定する新たな評価基準を設計し、CBIR に適用することにより、提案手法の有効性を論じる。

最後に、第 5 章では、本論文のまとめと今後の課題について述べる。

第 2 章

パノラミック検索モデルに基づくインタラクティブなマルチメディア検索手法

2.1 はじめに

本章では，人間と検索システムとの間に存在する Semantic Gap の解消に向けて，人間を積極的に検索ループの中に取り込むことを考え，インタラクティブなマルチメディア・コンテンツの検索モデル“パノラミック検索モデル”とそれに基づくマルチメディア検索手法“パノラミック検索手法”を提案する．また，Web ページ，映像などのマルチメディア・コンテンツの検索に適用することによって，提案モデルおよび手法の有効性を示す．

2.2 パノラミック検索モデル

提案するパノラミック検索モデルは，画像やテキストなどから構成されるマルチメディア・コンテンツを対象とし，

- 検索結果のコンテンツ集合と検索過程の仮想三次元空間による視覚化
- 画像特徴による視覚的な検索やテキスト特徴による意味的な検索など，マルチメディア・コンテンツを構成する各メディアの特性および各メディア間の関連性を活用し，各メディアに適した検索手段を有効に組み合わせたクロスメディア検索

を特徴とする．

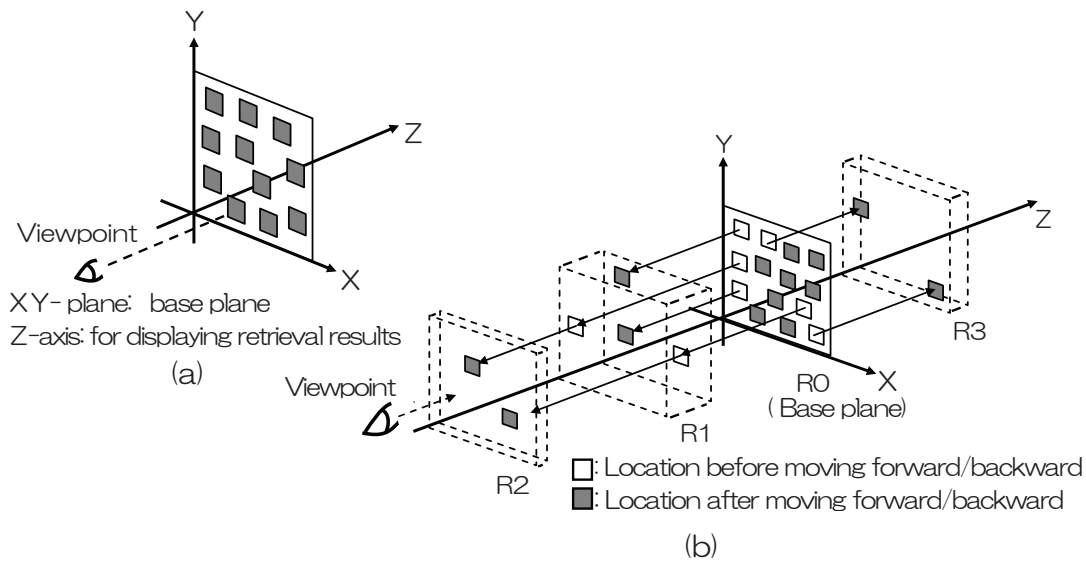


図 2.1 仮想三次元空間と検索過程の視覚化

本モデルは広くマルチメディア・コンテンツ全般に適用可能であるが、ここでは、検索対象となるコンテンツ集合として、その集合の要素が画像とテキストとから構成されるマルチメディア・コンテンツ（各々の要素が画像とテキストの対であるもの）を例にとって論じる。

2.2.1 コンテンツ集合と検索過程の視覚化

パノラミック検索モデルでは、図 2.1(a) に示す仮想三次元空間上で検索が進められる。

まず、検索対象のマルチメディア・コンテンツ集合が蓄積されたデータベースから、コンテンツ集合を検索するために、画像、あるいは、キーワードなどのテキストを検索クエリとして与える。検索クエリが画像の場合には、色や形状などの画像特徴を抽出し、それが類似した画像を含むコンテンツをデータベースから取り出す。一方、検索クエリがテキストの場合には、与えられたキーワードを含むコンテンツをデータベースから取り出す。この検索を初期検索と呼ぶ。初期検索により得られたコンテンツ集合は、図 2.1(a) に示す XY 平面（一覧表示平面と呼ぶ）上に灰色の小矩形のように配置される。このときの配置方法は、画像特徴あるいはテキスト特徴に基づく互いの類似関係を利用して、類似したコンテンツを近くに配置する。コンテンツの表示には、例えば、そのコンテンツに含まれる画像を用いる。

ユーザは一覧表示平面を垂線（Z 軸）方向から眺める。一覧表示平面ではコンテンツ集合が互いの類似関係に基づいて配置されているため、全体的に眺めることでコンテンツ集

合の大局的な構造が把握できる．また，特定の領域だけに注目することで，興味あるコンテンツとそれに類似するコンテンツを効率よく閲覧できる．このような一覧表示の方法は，所望のコンテンツを見出すための有効な手助けとなる．

しかし，表示されるコンテンツの数が多い場合には，さらに検索クエリを追加して所望のコンテンツを選別したり，不要なコンテンツを排除したりするといった絞込み検索が必要となる．

本モデルでは，絞込み検索の結果を一覧表示平面の垂線（Z 軸）方向に移動して表示することで，検索過程を視覚化する．この移動方向の軸（Z 軸）を検索過程表示軸と呼ぶ．図 2.1(b) に検索過程の視覚化方法を模式図で示す．

初期検索結果の一覧表示平面（R0）に対して，所望のコンテンツを選別するための検索クエリを与えると，選別した結果のコンテンツ集合がユーザの手前方向に移動して表示される（R1）．このとき，与えられた検索クエリとの類似度が高いものほどユーザに近い位置に表示される．さらに検索クエリを追加して選別した場合には，選別されたコンテンツがさらに手前に表示される（R2）．また，一覧表示平面（R0）に対して不要なコンテンツを排除する検索クエリを与えると，合致したコンテンツは検索クエリとの類似度が高いものほど奥行き方向に移動して表示される（R3）．また，これらの操作は任意の時点までバックトラックすることで，元の状態に戻すことができる．

このような絞込み検索の導入とその視覚化により，各段階での検索の結果得られるコンテンツ間の関係を容易に把握できる．

2.2.2 クロスメディア検索

検索対象のマルチメディア・コンテンツが画像とテキストとで構成される場合，初期検索や絞込み検索の手段として，色や形状などの画像特徴による検索とキーワードなどのテキスト特徴による検索とが考えられる．また，一覧表示平面への配置の手段についても，画像特徴間の類似関係による方法とテキスト特徴間の類似関係による方法とが考えられる．

初期検索や絞込み検索，一覧表示平面の配置のいずれにおいても，ユーザが検索対象の視覚的な特徴（例えば，対象が商品情報であれば，色や形状などのデザイン）に着目する場合には，画像特徴による検索・配置が，検索対象の意味的な内容（例えば，価格やブランド名など）に着目する場合には，テキスト特徴による検索・配置が有効である．

本モデルでは，インタラクティブな検索過程の各段階で，これらの検索対象の構成要素である画像，テキストといった各メディアの特性を，ユーザの目的に応じて適宜使い分け

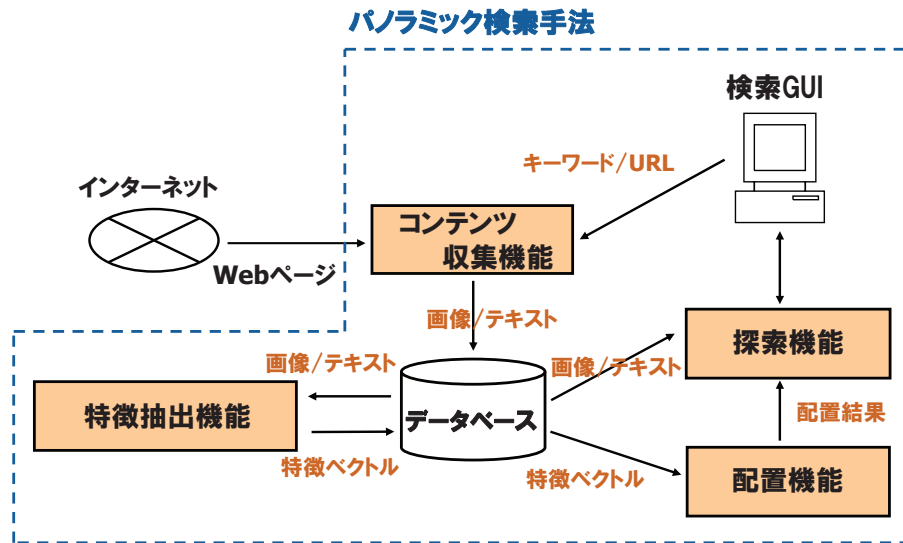


図 2.2 パノラミック検索手法の構成

ながら検索を進めていくこと，すなわち，クロスメディア検索を実現できる．

2.3 パノラミック検索手法

ここでは，パノラミック検索モデルを具現化したパノラミック検索手法の概要を述べる．

図 2.2 の破線部分に，パノラミック検索手法の構成を示す．

本手法は，

- 検索対象となるコンテンツ（画像およびテキスト）を収集するコンテンツ収集機能
- 収集した画像／テキストを管理するデータベース
- 収集した画像／テキストから特徴量（特徴ベクトル）を抽出する特徴抽出機能
- 抽出した特徴量に基づいて仮想三次元空間の一覧表示平面に画像／テキストを配置する配置機能
- 配置されたコンテンツ集合の絞込み検索を実行する探索機能
- 表示およびユーザとのインタラクションのための検索 GUI

から構成される．

なお，本手法では，一般的に処理時間のかかるコンテンツ収集機能および特徴抽出機能については，ユーザの利用に先立って事前に処理しておき，その他の機能については，ユーザの検索要求に基づいてリアルタイムに処理する構成を採っている．

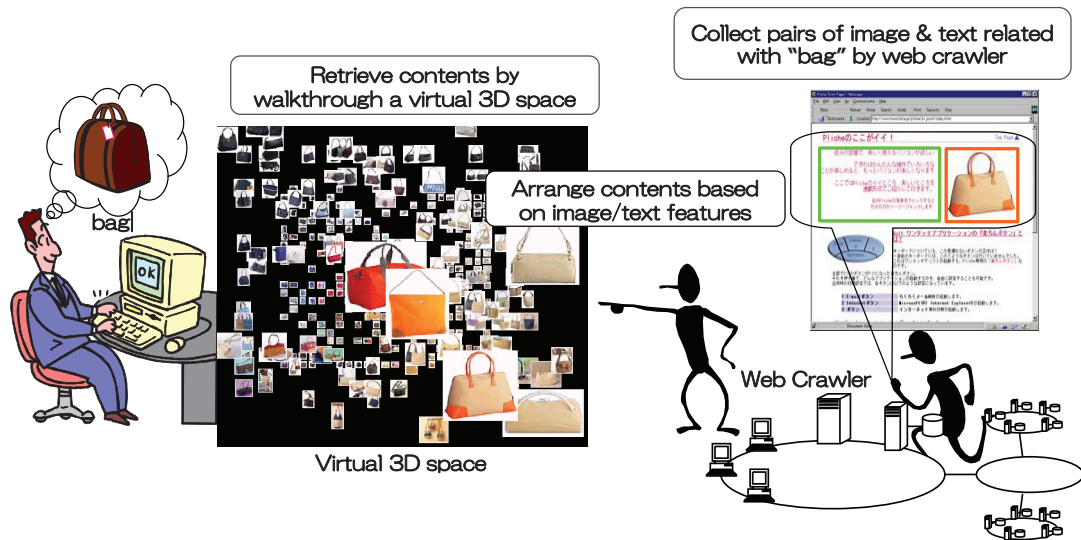


図 2.3 Web 検索システム

各機能の詳細については、次節の Web 検索への適用例を通じて具体的に述べる。

2.4 Web 検索システム

図 2.3 に、パノラミック検索手法を適用した Web 検索システムの概要を示す。この場合、パノラミック検索手法のコンテンツ収集機能の収集対象は、WWW (World Wide Web) 上のコンテンツとなる。

一般に、Web ページ中の画像周辺には、その画像と関連するテキスト、例えば、その画像の意味内容を説明するテキストが存在する。このようなテキストを画像からみて“関連テキスト”と呼ぶ。

図 2.4(a) に、関連テキストの例を示す。

本システムでは、Web ページ中の画像を検索するために、画像自体の特徴とともに、画像の意味内容を表すと考えられる関連テキストの特徴も利用する。

2.4.1 コンテンツ収集

コンテンツ収集機能により、WWW 上の画像とその関連テキストが収集される。二通りの方法で収集範囲を指定できる。一つは、URL (Uniform Resource Locator) を指定する方法であり、指定された URL を起点に指定された数だけ Web ページのリンクをた

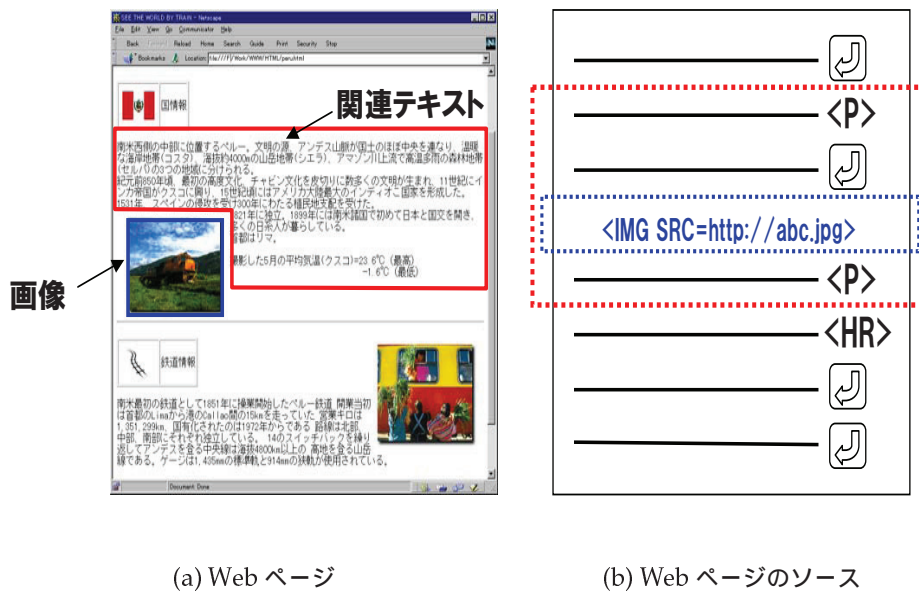


図 2.4 画像と関連テキスト

どり，Web ページ内にある画像とその関連テキストを収集する．もう一つは，収集用キーワードを指定する方法であり，入力されたキーワードをテキスト検索エンジンに渡し，その結果得られる URL のリストをたどって，Web ページ内にある画像とその関連テキストを収集する．

このとき，収集した画像の関連テキストは，図 2.4 に示すように，HTML (Hyper Text Markup Language) ファイルを解析して取得される．画像とテキストとの関連度を，

- 画像と近接する位置にあるテキストは，画像と内容的に関連していることが多い．
- HTML のタグ構造から画像とテキストとの関連度を推測できる．

という二つの観点から定義し，関連度の高いテキストを関連テキストとして画像とともにデータベースに保存する．

2.4.2 特徴抽出

特徴抽出機能は，画像およびテキストから特徴量（特徴ベクトル）を抽出する．これらの特徴量は，仮想三次元空間の一覧表示平面上への配置に用いられる．

CBIR の研究分野では，さまざまな画像特徴が提案されているが [4]，[18]，本システムでは，画像の色特徴として，HSI 色ヒストグラム特徴，HSI 色レイアウト特徴 [19] など

を，画像の形状特徴としては，Wavelet 変換係数特徴 [20]，[21] などを，また，テキスト特徴としては，キーワードやベクトル空間モデルに基づく単語頻度特徴 [22]，[23] などを用いる．

2.4.3 配置

配置機能は，初期検索によりデータベースから取り出したコンテンツ集合の画像 / テキスト特徴量の類似性に基づいて，画像集合を仮想三次元空間の一覧表示平面に配置する．互いの特徴ベクトル間の距離が小さい画像が，一覧表示平面上でも近い距離に配置される．

初期検索の検索クエリとしてテキストを与える場合には，そのテキストによりデータベースから取り出すコンテンツ集合を絞り込むことができるが，画像を与える場合には，一般に，データベース内のすべての画像との画像特徴量間の類似度計算が必要となり，時間がかかる．ここでは，この問題に対処するために，多次元インデクス方式 [24] を実装し，類似度計算の高速化を図っている．

配置方式として，Kohonen の提案した教師なしの競合学習モデルである自己組織化マップ (Self-Organizing Map, SOM) を用いる [25]．SOM では，データの分布の把握を容易にするために，高次元の特徴ベクトル空間を低次元の空間に写像する．その際，高次元空間における距離が近いデータほど，低次元空間においても近くに配置される．ここでは，低次元の空間として二次元空間（仮想三次元空間の一覧表示平面）を採用し，画像を格子状に配置する．また，画像の色や形状の特徴やテキストの単語頻度特徴など，コンテンツ特徴の種別ごとに配置情報を生成するので，ユーザは目的に応じてコンテンツ特徴の種別を適宜切り替えて表示させることができる．

図 2.5 は，電機メーカーの URL を指定して収集したコンテンツ集合を画像の色特徴 (HSI 色ヒストグラム特徴) で配置した例を示す．同系色のノート PC の画像が図の中心付近に集まって配置されている．

また，図 2.6 は，映画紹介サイトの URL を指定して収集したコンテンツ集合をコンテンツ中のテキストの特徴 (単語頻度特徴) に基づいて配置し，その配置位置にコンテンツ中の画像を表示した結果を示す．

テキスト特徴による配置の場合には，画像特徴による配置とは異なり，画像の視覚的な類似性を利用してコンテンツの類似性を把握することが困難である．本システムでは，ナビゲーションのためのランドマークとして，テキストから自動的に抽出した tf・idf 法 [26] に基づく重要単語である“キーワードラベル”を表示する．この機能を用いること

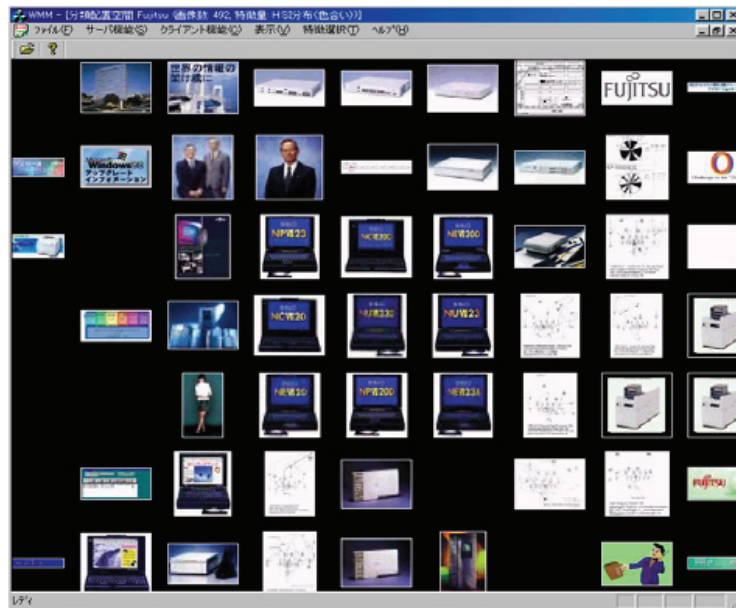


図 2.5 画像の色特徴による一覧表示平面の表示例

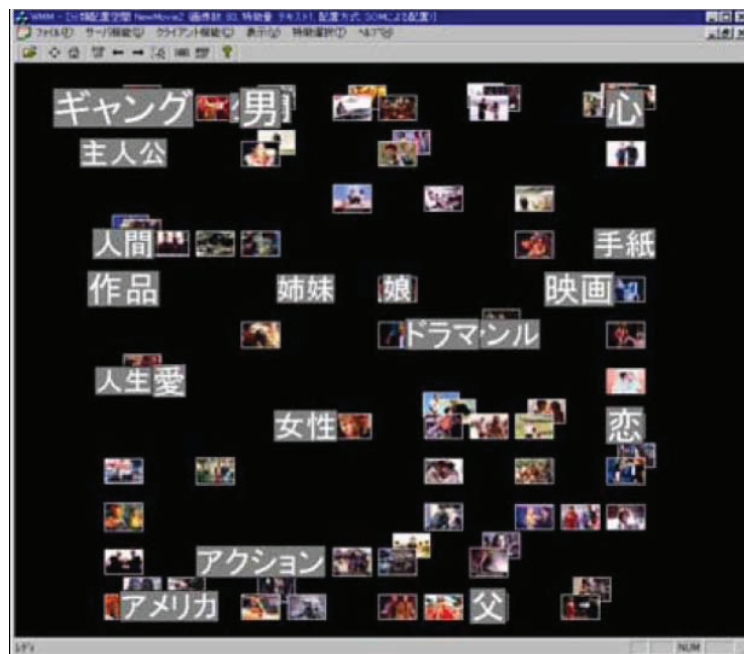


図 2.6 テキスト特徴による一覧表示平面の表示例

で、ユーザはキーワードラベルを参照しながら、どの辺りにどのようなコンテンツがあるのかを容易に把握でき、興味ある部分に近づいてその近辺の関連したコンテンツだけに集中して閲覧することができる [23]。

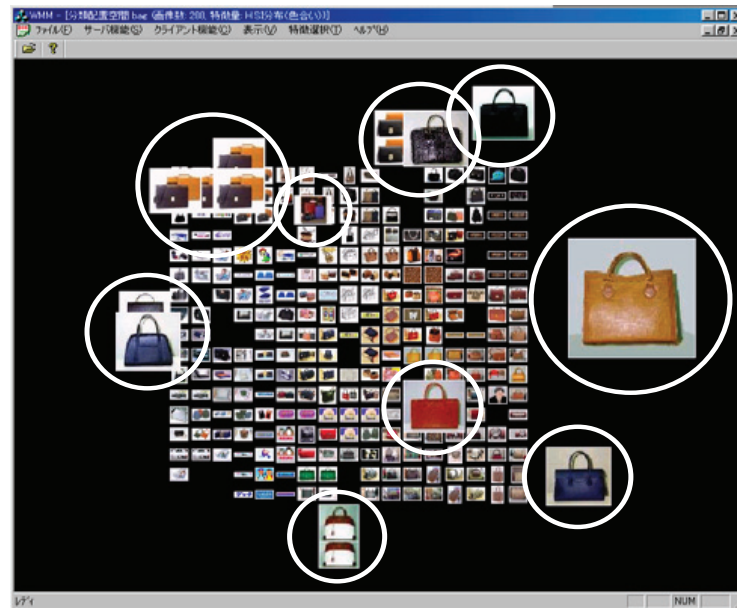


図 2.7 選別による絞り込み検索結果の表示例

仮想三次元空間の一覧表示平面に配置されたコンテンツ集合は、図 2.5、図 2.6 の表示例のように、ユーザに提示される。提示直後の視点は、コンテンツ集合全体を概観できる位置に設定され、その後、ユーザは興味ある画像付近に視点位置を移動（ウォークスルー）して詳細を確認したり、また、画像を指定することにより画像の関連テキストを表示できる。

2.4.4 探索

一覧表示平面に表示されたコンテンツ集合を対象とした探索機能として、上述の絞り込み検索機能と類似コンテンツリスト表示機能とを備える。これらの探索機能は、コンテンツ集合の画像あるいはテキストの特徴を用いて実行できる。

絞り込み検索機能は、一覧表示平面に表示されたコンテンツに対し、ユーザがキーワードなどの検索クエリを与えることにより実行される。絞り込み検索には、ユーザが与えた検索クエリに合致するコンテンツを選別する方法と合致するコンテンツを排除する方法とがある。

図 2.7 は、検索クエリに合致するコンテンツを選別する絞り込み検索結果の表示例である。この例は、バッグの商品情報の検索を想定したものであり、キーワード“バッグ”を指定して収集したコンテンツ集合を画像の形状特徴に基づいて一覧表示平面に配置している。ここで、ビジネス用のバッグの情報のみを選別するために、キーワード“ビジネス



図 2.8 排除による絞り込み検索結果の表示例

バッグ”で絞り込み検索を実行した結果，関連テキストに“ビジネスバッグ”を含む画像が手前に移動して大きく表示される

図 2.7 では，白丸で囲んだ画像が選別結果であり，絞り込み前のコンテンツ集合との位置関係が一覧可能となっている様子が分かる．

また逆に，不要なコンテンツを排除する検索クエリを与えると，与えられた検索クエリと合致するコンテンツが画面の奥に移動し，小さく表示される．図 2.8 は，不要なコンテンツを排除する絞り込み検索結果の表示例である．この例は，バッグ（鞆）を想定して入力したキーワード“バッグ”により収集した画像の中に自動車のエアバッグ画像が含まれていたため，キーワード“エアバッグ”によって排除した例である．

図 2.8 では，楕円で囲まれた領域に，奥に移動したエアバッグ画像が表示されている．

このように，不要なコンテンツを画面上から消去するのではなく，画面の奥に目立たなく表示しておくことで，検索過程の履歴情報として利用できる．

また，コンテンツ集合の移動の過程をアニメーションで連続的に表示することにより，移動するコンテンツを容易に把握することができる．なお，キーワードなどのテキスト特徴の代わりに画像特徴を用いる場合には，検索キーとする画像と色や形状などの画像特徴を指定することによって，画像特徴による絞り込み検索を実現できる．

検索対象がある程度絞り込まれた段階では，ある特定のコンテンツに類似したものを類似度順に閲覧するのが効率的である．類似コンテンツリスト表示機能によって，ユーザは

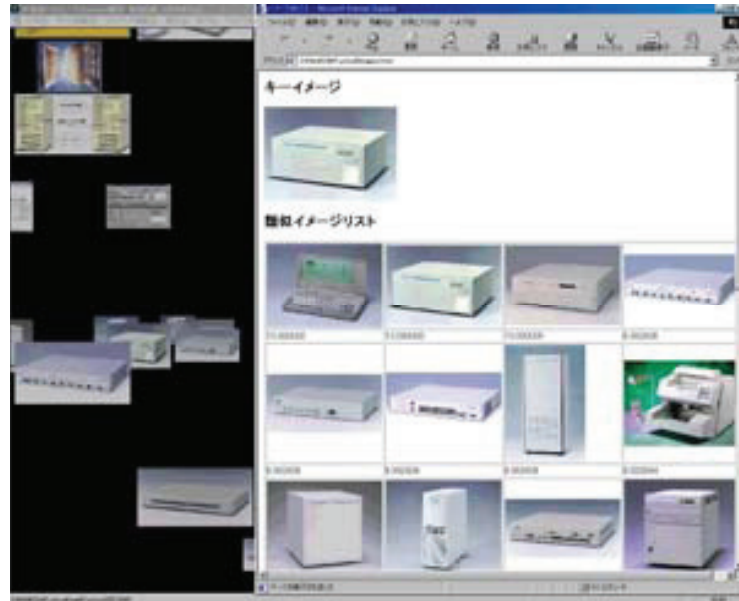


図 2.9 画像の形状特徴による類似コンテンツリストの表示例

検索キーとする特定のコンテンツと画像／テキスト特徴を指定すると，検索キーのコンテンツをリストの先頭とし，指定された特徴に基づいた類似度順に類似コンテンツを並べて表示できる．

図 2.9 に，画像の形状特徴（Wavelet 変換係数特徴）に基づく類似コンテンツリストの表示例を示す．

最終的に，ユーザが目的に合った所望のコンテンツを見出し，それを指定すると，元々そのコンテンツが存在した Web ページが表示される．

なお，パノラミック検索手法では，仮想三次元空間において大量の画像を高速に表示するために，いわゆる LOD (Level of Details, LOD) 表示を採用している．図 2.10 に示すように，個々のコンテンツに対して解像度の異なるいくつかのサムネイル画像を準備しておき，ユーザの視点の移動に応じて，画面に大きく表示する場合には高解像度の画像を，小さく表示する場合には低解像度の画像を使用することにより，ファイルの読み込み時間および描画時間を短縮し，インタラクティブな検索にとって重要な要素である高速表示を実現している．これによって，1GHz の Intel Pentium III 搭載 PC で，数千枚の画像をスムーズに表示，閲覧することができるが，画像の一覧性の良さを利用するという観点からは，一度に最大 5～6 百枚程度を表示，閲覧するのが適切であると考えている．

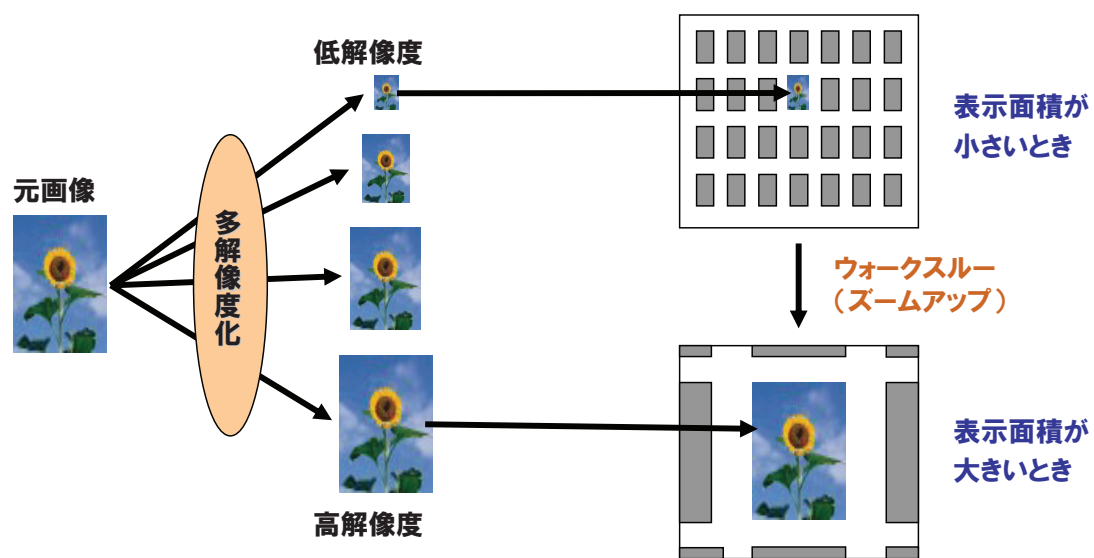


図 2.10 LOD 表示

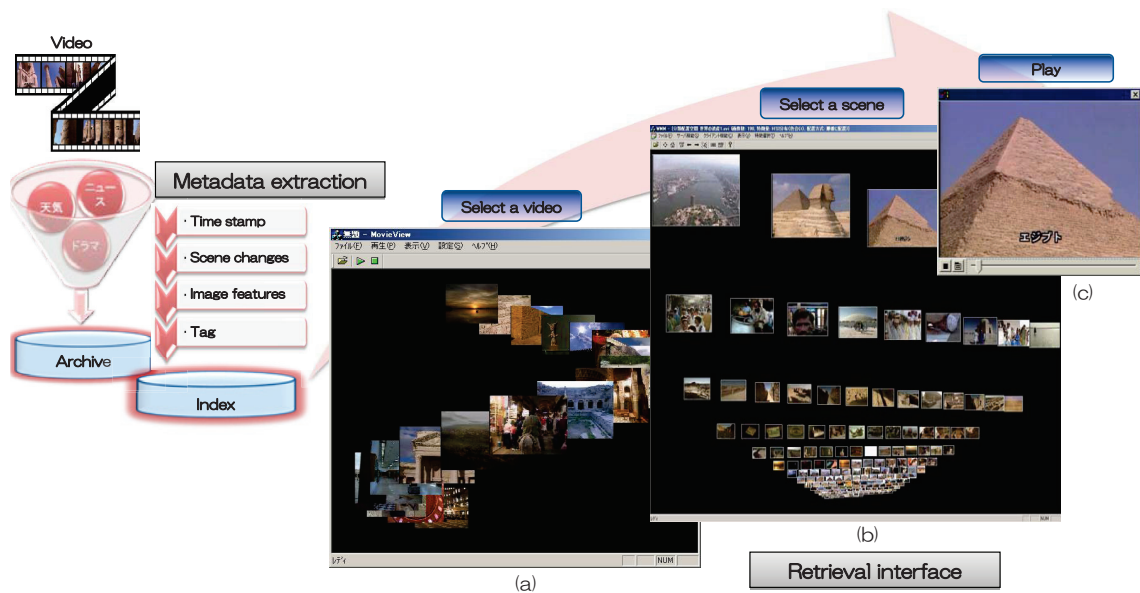


図 2.11 映像検索システム

2.5 映像検索システム

大量のコンテンツをユーザに把握しやすく提示するというパノラミック検索手法の枠組みは、そのまま映像の検索にも適用できる。映像検索は、大量の映像コンテンツ集合の中から所望のコンテンツを検索する映像コンテンツ検索と、個々の映像コンテンツからある特定のシーンを検索する映像シーン検索とに分けられる。

図 2.11 に、TV 映像などの編集済み映像を対象としたパノラミック検索手法による映像検索システムの概要を示す。

本システムは、大量の映像コンテンツ集合を同時に再生しながらユーザに提示する映像コンテンツ検索機能（図 2.11(a)）と、ユーザにより選択された一つの映像コンテンツから自動抽出したシーン代表画像（シーン変化点直後のフレーム画像）の集合を一覧表示し、特定のシーンをインタラクティブに検索する映像シーン検索機能（図 2.11(b)）とから成る。最終的には、ユーザが所望のシーン代表画像を検索、指定することにより、そこから映像シーンが再生される（図 2.11(c)）。

2.5.1 映像コンテンツ検索

映像コンテンツ検索機能は、大量の映像コンテンツ集合をユーザに分かりやすく同時に提示し、ユーザの目的に合った所望のコンテンツを検索させる機能である。ここでは、図 2.11(a) に示すように、螺旋状に複数の映像コンテンツを配置し、映像を再生しながら流していく螺旋状の映像の同時再生方法を実装した。映像コンテンツが自動的に再生されながら流れていくので、ユーザは多くの映像コンテンツに対しても、画面の中央付近にきた映像に着目するだけで、順次、映像コンテンツの内容を把握し、確認できる。

なお、この映像コンテンツの同時再生は、映像検索システムにおいて最も負荷のかかる処理であるが、30 本程度の映像コンテンツ（フレームサイズ：160 × 120）であれば、1GHz の Intel Pentium III 搭載 PC で十分滑らかに再生できる。

2.5.2 映像シーン検索

映像シーン検索機能では、ある一つの映像コンテンツから自動的にシーン変化点を検出し、一つの映像コンテンツをシーン代表画像の集合として表現する。また、シーン代表画像と関連付けられたクローズドキャプションなどの関連テキストが存在するならば、一つの映像コンテンツをシーン代表画像と関連テキストの対の集合として表現することもできる。したがって、映像シーン検索は、シーン代表画像（あるいはシーン代表画像と関連テキストの対）の集合の中から所望のシーン代表画像を検索する問題と捉えることができ、前節で述べたパノラミック検索手法のさまざまな機能が映像シーン検索においてもそのまま適用できる。図 2.11(b) に示すように、ユーザは仮想三次元空間に表示されたシーン代表画像群の中から所望の画像を検索することによって、ある特定のシーンを再生することができる（図 2.11(c)）。

図 2.12 および図 2.13 に、それぞれ、シーン代表画像の色特徴（HSI 色ヒストグラム特徴）による表示例とテキスト特徴（単語頻度特徴）による表示例を示す。

また、時間軸を持つ映像特有の表示機能として、シーン代表画像のフィルム状の一覧表示も実現した（図 2.11(b)）。これは映像の時間軸を空間的に展開するもので、シーン代表画像をこのように配置することにより、ユーザは映像を再生 / 早送りして閲覧することなしに、その映像がどのようなシーンからどのような順序で構成されているかを一目で把握できる。

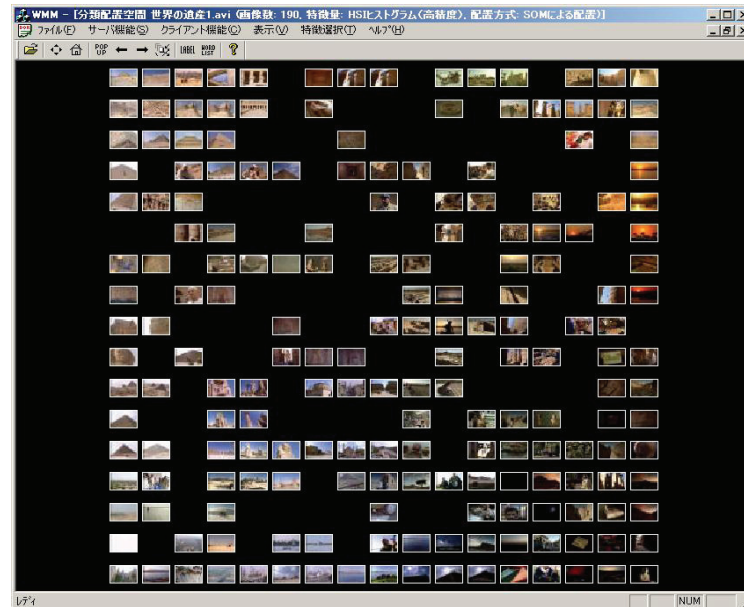


図 2.12 シーン代表画像の色特徴による表示例



図 2.13 シーン代表画像のテキスト特徴による表示例

表 2.1 パノラミック検索手法による検索対象の表現図式

応用	画像 (or 3D モデルなど)	関連テキスト
Web 検索	画像	関連テキスト
映像検索	シーン代表画像	クローズドキャプションなど
デジカメ画像検索	デジカメ画像	撮影時刻など
デジタル文書検索	スライドのサムネイル画像 (PowerPoint 文書) など	スライド中のテキスト (PowerPoint 文書) など
図面検索	図面	設計者, 作成日, 図番など
3D-CAD モデル検索	3D-CAD モデル	設計者, 作成日, 図番など

2.6 その他の応用

上述したように, パノラミック検索手法では, 検索対象を“画像と関連テキスト”の対として捉えて表現する. Web 検索では, Web ページを“画像とその周辺の関連テキスト”の対として, 映像検索では, 映像シーンを“シーン代表画像とクローズドキャプションなどの関連テキスト”の対として表現した.

検索対象とするマルチメディア・コンテンツにこの“画像と関連テキスト”の対による表現図式をあてはめられるならば, パノラミック検索手法の枠組みがそのまま適用できる. Web 検索, 映像検索以外にも, デジカメ画像 [27], デジタル文書 [28], 図面 [29], 3D-CAD モデル [30], [31] など, さまざまなマルチメディア・コンテンツにパノラミック検索手法を適用している.

デジカメ画像の場合には, “デジカメ画像と撮影時刻”, デジタル文書(例えば, Microsoft PowerPoint 文書)の場合には, “各スライドのサムネイル画像とスライド中のテキスト”, 図面あるいは 3D-CAD モデルの場合には, “図面あるいは 3D-CAD モデルと設計者, 作成日, 図番などの付随情報”の対として検索対象を表現し, パノラミック検索手法による効率的な検索を実現できる.

表 2.1 に, 検索対象の表現図式の要約を示す.

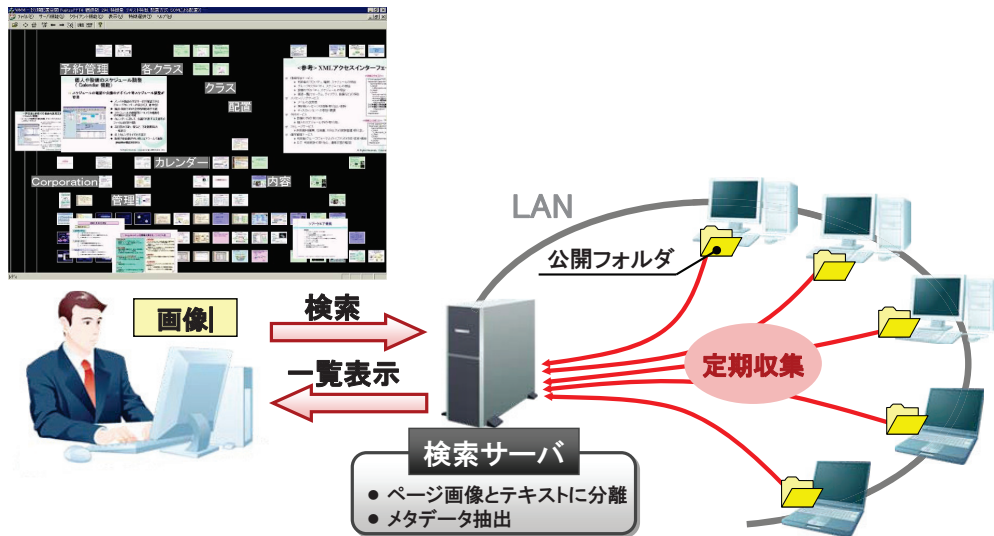


図 2.15 デジタル文書検索システム

収集されて検索サーバに自動的に登録され、グループ内で共有される。ユーザは、パノラミック検索手法のさまざまな機能を使って、検索サーバに登録されたデジタル文書を効率的に検索することができる。

2.6.3 図面検索システム

機械メーカーなどでは、設計段階において設計者が新規に図面を作成する際、大量に蓄積された既存図面の中から参考となる図面を検索し、図面に記載された設計上のノウハウを参照したり、図面の一部を流用したりすることにより、図面の作成コストだけでなく、部品の調達コストや保守管理コストを削減する努力がなされている。

このようなコスト削減を支援するために、大量の既存図面の中から所望の図面を検索する図面検索システムを開発した。図 2.16 に、組立図面を対象とした図面検索システムの概要を示す。

本システムは、予め図面を部品毎に自動的に分割しておき、設計者がクエリ画像を与えると、そのクエリ画像と形状が類似した部品が記載された図面とその図面上の該当部品の位置を明示した形（図 2.16 の緑色の矩形）で検索結果を類似度順に提示することができる。

なお、線画である図面に記載された部品画像に適した特徴として、Zernike モーメント特徴 [32]、Grid Fourier 特徴 [29]、[33]、エッジ方向ヒストグラム特徴 [34] などを用いている。

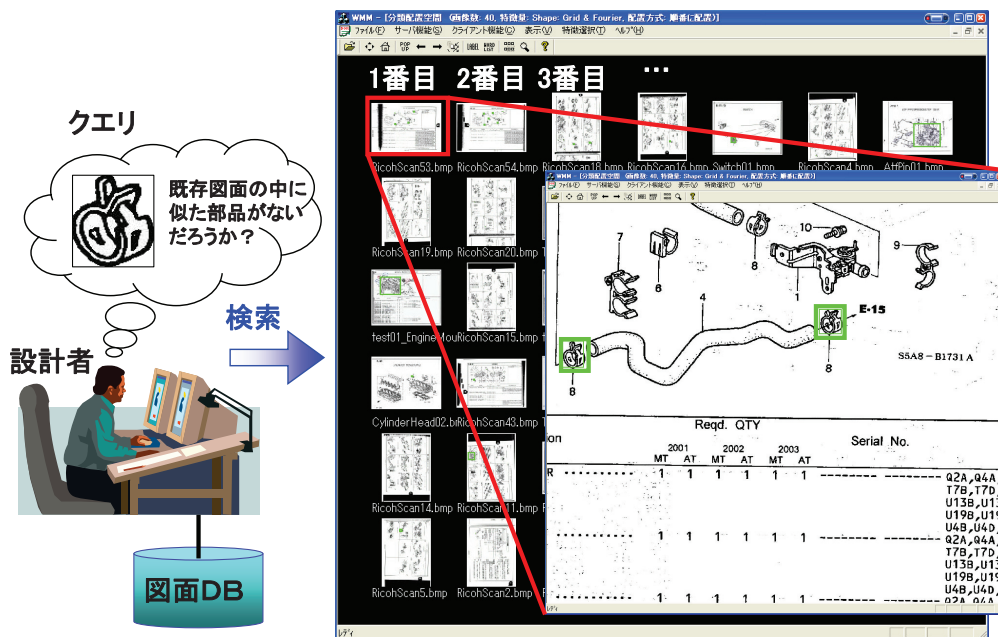


図 2.16 図面検索システム

2.6.4 3D-CAD モデル検索システム

CAD, 電子商取引, エンターテインメントなど, 多くの分野で三次元形状モデルが広く使用されるようになってきており, 大量の三次元形状モデルを格納したデータベースの中から所望の三次元形状モデルを効率的に検索するシステムが必要とされている. 上述した図面検索システムと同様に, 特に, CAD 分野においては, 設計コストなどの削減を実現するための 3D-CAD モデル検索システムに対するニーズは高い.

図 2.17 に, このニーズに応えるために開発した 3D-CAD モデル検索システムの概要を示す.

3D-CAD モデルの場合には, 検索対象を“3D-CAD モデルと付随情報(設計者, 作成日, 図番など)”の対で表現することによって, パノラミック検索手法の枠組みはそのまま適用できるが, Web 検索や映像検索など“(二次元)画像”を検索対象とする場合と異なり, 三次元形状モデル特有の機能を追加する必要がある.

一つ目は, 三次元形状に対する特徴抽出機能の追加である. CAD 分野を中心にさまざまな三次元形状特徴量が提案されているが [31], [35], [36], [37], ここでは, Shape Distribution 特徴 [36] を用いている.

二つ目は, 2.4.4 節で述べた LOD 表示機能の拡張である. 複雑な 3D-CAD モデルを描

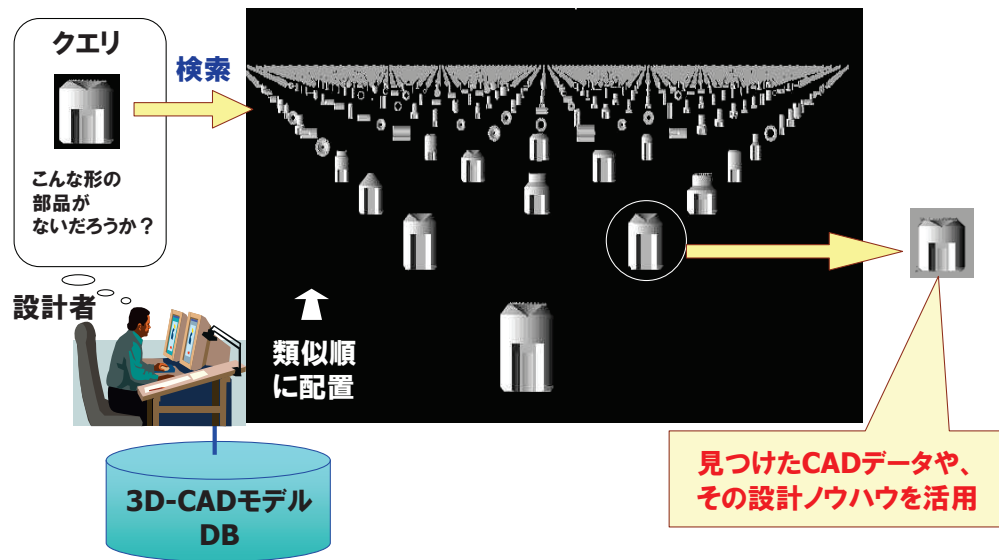


図 2.17 3D-CAD モデル検索システム

画する場合、ポリゴンをレンダリングするのではなく、予めポリゴンをレンダリングした結果の画像を描画すれば、高速であり、その描画時間は、3D-CAD モデルの複雑さには依存しない。ここでは、大量の 3D-CAD モデルを高速にレンダリングするために、ユーザの視点と 3D-CAD モデルとの間の距離に応じて、“ポリゴン”，ユーザの視点方向からのレンダリング“画像”，モデルの存在のみを表す“点”を適切に切り替えて描画するハイブリッド・レンダリング機能を実装した。なお、図 2.17 に示すように、類似コンテンツリスト機能の一つとして、3D-CAD モデルを鳥瞰図として表示する機能も追加した。

三つ目は、配置機能の拡張である。例えば、Shape Distribution 特徴に基づいて 3D-CAD モデルを仮想三次元空間の一覧表示平面に配置する場合、個々の 3D-CAD モデルをどのようなポーズ（姿勢）で配置すべきかという問題が生じる。図 2.18 に、SOM を用いた 3D-CAD モデルの Shape Distribution 特徴による表示例を示す。図中の領域 A には、細長い形状のモデルが、領域 B には、中央に穴が開いたモデルが集まって配置されている。しかし、ユーザが形状を把握するのに適していないポーズの 3D-CAD モデルがいくつかある。例えば、領域 C の三つのモデルは、互いに類似した形状を持つが、異なるポーズをとっているため、その把握が難しくなっている。また、領域 D の二つのモデルは、中央に穴を持つが、現在のポーズではその穴を見ることはできない。この 3D-CAD モデル特有の問題を解決するために、配置機能を拡張し、ポーズ・アラインメント機能を実装した。

以下では、このハイブリッド・レンダリング機能とポーズ・アラインメント機能について、その概要を述べる。

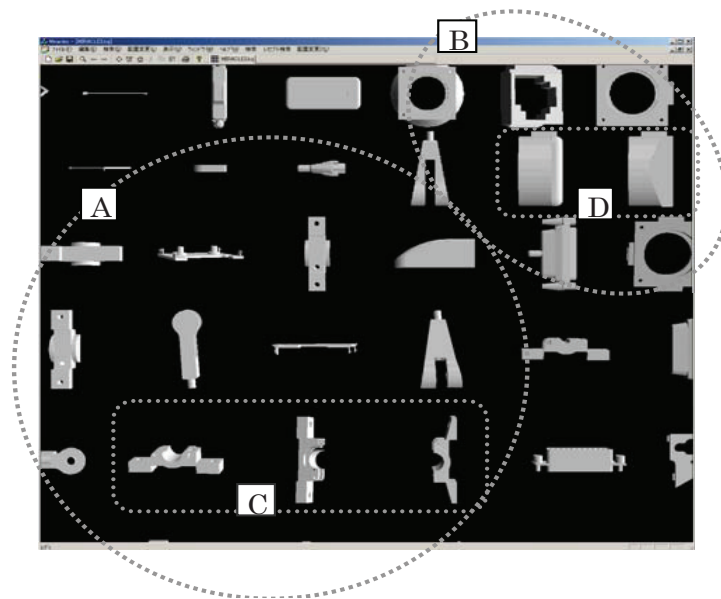


図 2.18 3D-CAD モデルの Shape Distribution 特徴による表示例

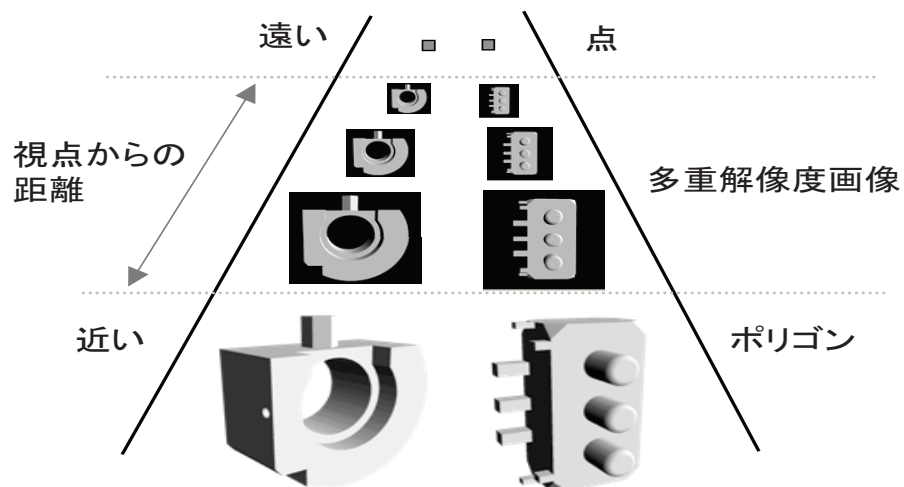


図 2.19 ハイブリッド・レンダリング

ハイブリッド・レンダリング

ハイブリッド・レンダリング機能では、図 2.19 に示すように、ユーザの視点と 3D-CAD モデルとの間の距離に応じて、3D-CAD モデルの表現を変える。3D-CAD モデルがユーザの視点から遠くにあれば、点で、近くにあれば、ポリゴンで、その中間の範囲にあれば、その距離に応じた多重解像度画像で描画される。

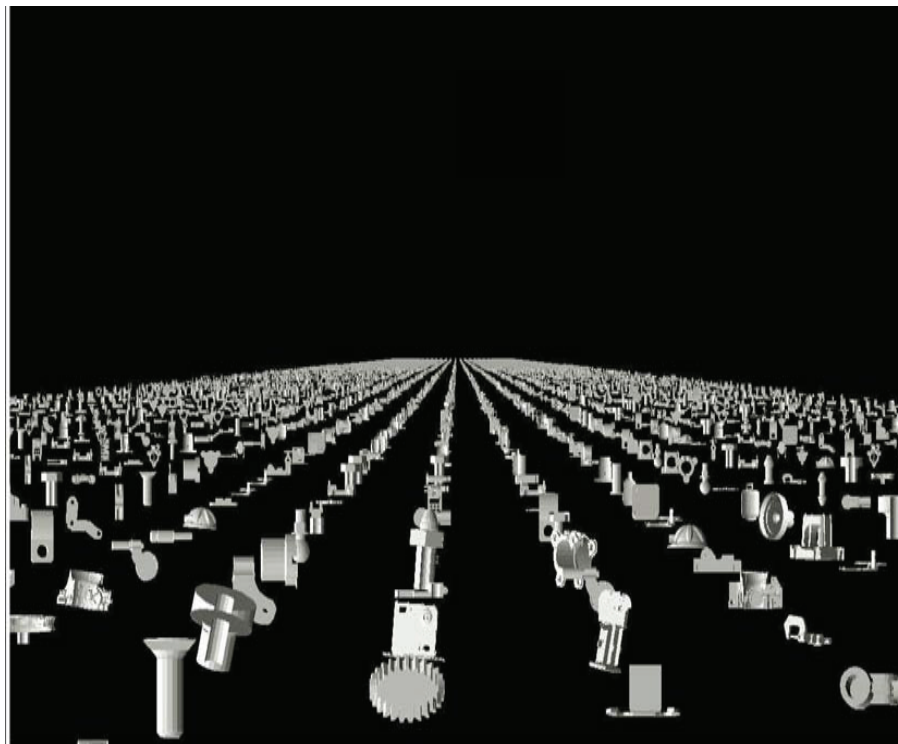


図 2.20 ハイブリッド・レンダリングによる 3D-CAD モデルの表示例

図 2.20 に、ハイブリッド・レンダリングによる表示例を示す。レンダリングの計算コストを削減したことにより、1GB メモリ、1.2GHz の Intel Core Solo 搭載 PC で、10,000 個以上の 3D-CAD モデルをスムーズに描画でき、点、画像、ポリゴンの三つの表現の切り替わりも目立たないことを確認した。

ポーズ・アライメント

上述したように、3D-CAD モデルは視点方向によっては全く異なる形に見えるため、表示する際には、どのようなポーズで表示すべきかが重要となる。ここでは、大局的ポーズ・アライメントと局所的ポーズ・アライメントの二つの機能を開発した。これらの機能では、設計作業の一般性は損なわない範囲の制約として、図 2.21 に示すように、設計者は x 軸、 y 軸、 z 軸のいずれかの軸に沿って 3D-CAD モデルを作成することを仮定し、アライメントのための 3D-CAD モデルの回転の自由度を x 軸、 y 軸、 z 軸のいずれかを軸にした 90 度単位の回転に制限している。

大局的ポーズ・アライメント

大局的ポーズ・アライメントは、ユーザが 3D-CAD モデル全体を概観し、所望の

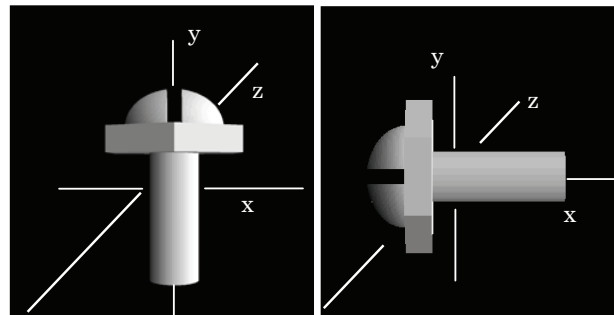


図 2.21 ポーズ・アラインメントの仮定

3D-CAD モデルが配置されている可能性が高い領域を把握しやすくすることを支援する機能である。そのために、個々の 3D-CAD モデルの長軸が上下方向を向くように、かつ、その方向から見えるポリゴンの数が一番多い方向に 3D-CAD モデルを回転させるというヒューリスティックな方法を採用している。

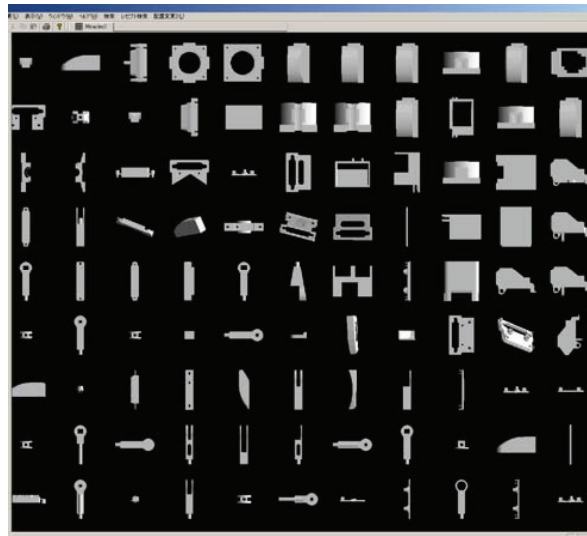
図 2.22(a) および図 2.22(b) に、大局的ポーズ・アラインメントを実行した前後の表示例を示す。図 2.22(b) では、すべての細長い形状の 3D-CAD モデルが上下方向に揃って配置され、また、図の右上部分では、中央に穴の開いたモデルが穴の部分を前方に向けるように配置されており、ユーザはどこに所望の 3D-CAD モデルが存在するかを把握しやすくなる。

局所的ポーズ・アラインメント

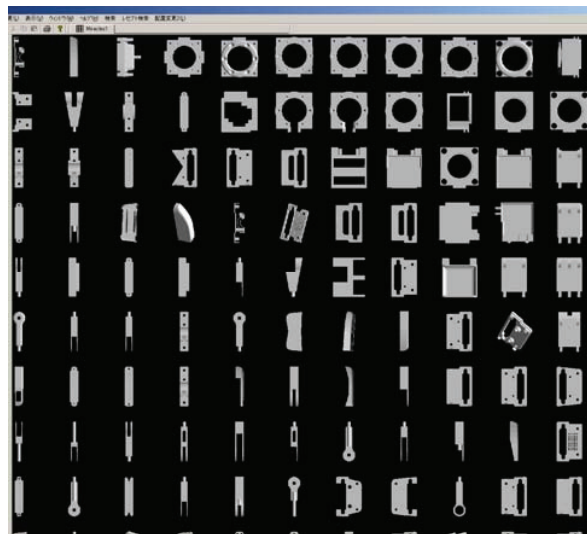
ユーザが、3D-CAD モデル全体を概観し、所望のモデルが存在する可能性の高い領域に近づくと、その領域には類似した形状の 3D-CAD モデルが集まっている。局所的ポーズ・アラインメントは、この類似した形状の 3D-CAD モデルの中から所望のモデルを見出すことを支援する機能である。

ここでは、図 2.23 に示すように、ユーザが類似した形状の 3D-CAD モデルの中から一つのモデル（図中、赤色の枠で囲ったモデル）を選択すると、その周辺の 3D-CAD モデルが、選択された 3D-CAD モデルと同じ方向に回転させられる。異なる方向から見比べたい場合には、ユーザのマウス操作に応じてすべての 3D-CAD モデルを同時に回転することができる。したがって、ユーザは各モデルを詳細に見比べて、所望の 3D-CAD モデルを容易に見出すことが可能となる。

具体的には、図 2.24 に示すように、選択された 3D-CAD モデルとその周辺のモデルから、それぞれ x 軸、 y 軸、 z 軸とその反対方向の 6 方向から見た場合の画像を生成し、選択されたモデルとその周辺のモデルがそれぞれの方向を向いた場合の対応する画像の特徴



(a) アラインメント前



(b) アラインメント後

図 2.22 大局的ポーズ・アラインメント

ベクトル (Wavelet 変換係数特徴) 間の類似度の総和を求め, その類似度の総和が最大となる方向に 3D-CAD モデルを回転させることによって, 局所的ポーズ・アラインメントを実現している。

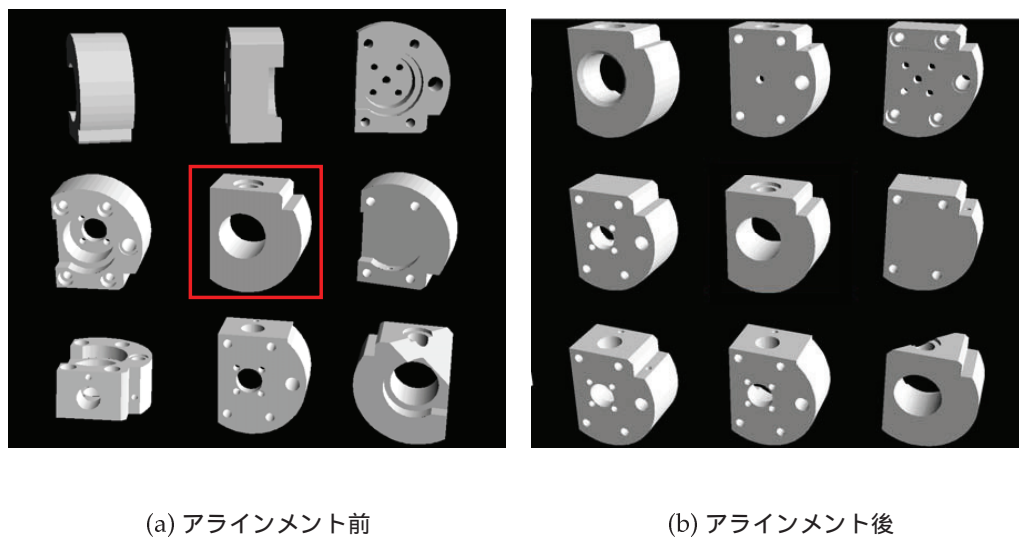


図 2.23 局所的ポーズ・アラインメント

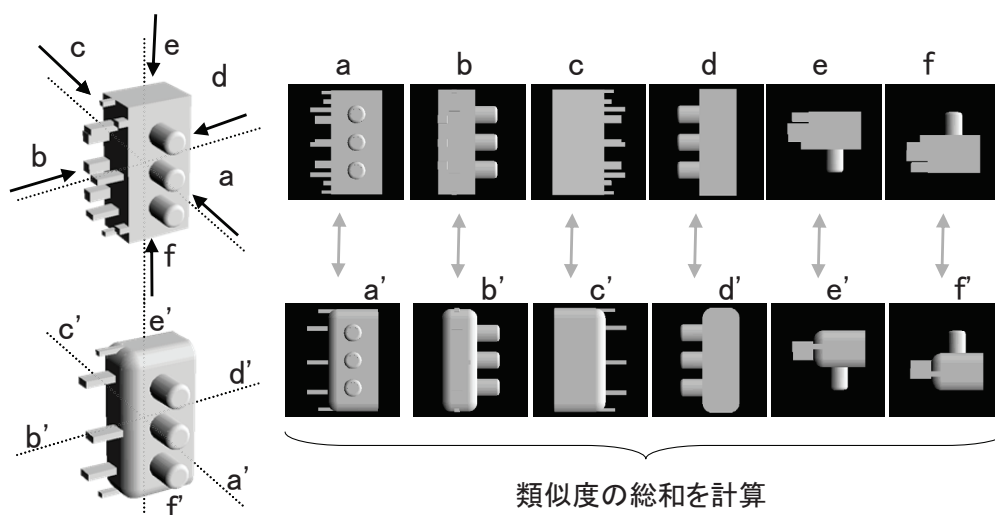


図 2.24 局所的ポーズ・アラインメントの類似度計算

2.7 関連研究

関連する研究あるいはシステムとして、武者らは、類似画像検索のための視覚化手法と GUI を開発した [6]。この手法は、検索対象画像群全体ではなく検索結果画像群の画像特徴量に主成分分析を適用し、検索結果画像群をその固有空間上の散布図として表示することにより、ユーザの類似画像検索を支援しようとするものである。

木本は、仮想三次元空間を用いたインタラクティブな画像検索モデルとシステムを提案した [7]。このシステムでは、仮想三次元空間（HSI 色空間の HSI を各軸とする直交三次元空間）の原点に検索キー画像を置く。検索キー画像と検索対象画像群との間の HSI 各色成分の分布の差を画像間の距離として定義し、その距離に基づいて検索結果画像群を直交三次元空間に配置・表示する。その画像群の中からユーザが検索意図にあった画像を次の検索キー画像として選択すると、その画像をまた原点に置き、検索結果画像群が再配置される。この過程を繰り返しながらユーザはインタラクティブに検索を進めることができる。

(株) データクラフトは、画像を画像特徴の類似性に基づいて仮想二次元 / 三次元空間に配置し、ナビゲーション GUI を通じて画像を検索する次世代画像クルージング技術「ImageCruiser」を提案している [8]。

本論文で提案するパノラミック検索手法は、HSI 色ヒストグラムなどの画像特徴に基づいて画像を仮想三次元空間に分類・配置し、インタラクティブに所望の画像を検索するという点では、これらの研究あるいはシステムと類似しているが、初期検索のみならず絞り込み検索も含め、すべてのインタラクティブな検索過程においてクロスメディア検索を実現しているという点で異なっている。

一方、クロスメディア検索という観点からの関連研究として、桑原らが提案したクロスメディア・メタサーチがある [38]。これは、複数の検索キーワードから成るクエリを分割し、テキスト検索エンジンや画像検索エンジンなど多様なメディア向けの既存の検索エンジンに渡して検索処理を行い、それらの検索結果を自動的に統合する手法である。

しかし、ユーザが与えられる検索クエリはキーワードだけであり、多様なメディアの特性を活かしてインタラクティブに切り替えながら検索を進める本論文のクロスメディア検索とは異なる。また、画像検索エンジンとしては、Google Image Search[39] など、キーワードベースのものを利用しており、CBIR は扱われていない。

2.8 実験

パノラミック検索手法の有効性を検証するために、デジタル文書検索システムおよび映像検索システムを用いた検索時間の比較評価実験を行った。

2.8.1 実験 1：デジタル文書の検索

パノラミック検索手法のクロスメディア検索および仮想三次元空間への配置、探索機能の有効性を検証するために、“画像と関連テキスト”の対で表現されたマルチメディア・コンテンツとして、デジタル文書を採り上げ、図 2.15 に示すパノラミック検索手法によるデジタル文書検索システム（以下、本節では、パノラミック検索システムと呼ぶ）と全文検索システムとの検索時間の比較実験を行った。いずれのシステムもクライアントサーバ方式で実現されており、Web ブラウザで利用できる。全文検索機能としては、Accela Bizsearch[40]を使用した。

実験データと実験方法

検索対象の文書として、213 件の Microsoft PowerPoint 文書（以下、PowerPoint 文書と呼ぶ）を使用した。スライド総数は 5,147 枚であった。各スライドは、そのサムネイル画像とスライド内のテキストの対で表現されており、文書単位では、タイトル、作成者、作成日時といったメタデータも付与されている。

実験では、パノラミック検索システムを利用して PowerPoint 文書をスライド単位で検索した場合と、全文検索システムを利用して PowerPoint 文書のタイトルと概要を一覧表示し、その中から文書を選択して開くことにより所望のスライドを検索する場合とで、図 / 画像や文書の転送・表示時間を含め、検索時間を計測、比較した。

また、パノラミック検索システムが初心者にも受け入れられるかを検証するためにも、全文検索システムには習熟しているがパノラミック検索システムの利用経験は少ない 10 名を被験者とした。検索タスクとして、デジタル文書検索における典型的な場面を想定し、以下の三つを被験者に課した。

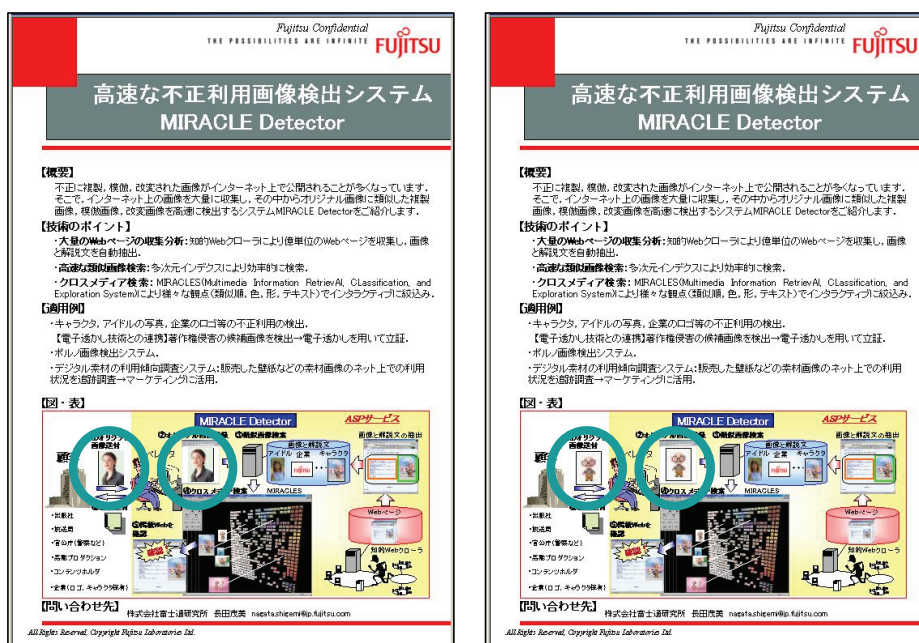


図 2.25 デジタル文書検索実験 (Task2)

Task1 キーワードのみで十分に絞り込める場面

「“イメージマイニング”のデモ操作手順を説明するスライドを探せ」とだけ被験者に指示した。クエリとして“イメージマイニング”、“デモ”の論理積で検索すれば、5件の文書、5枚のスライドに絞り込まれ、その5件についても、文書のタイトルから内容を容易に推定できる。

Task2 図 / 画像の一部が変更されたスライドが多数存在する場合

図 2.25(a) のスライドを被験者に提示し、「スライド下部のシステム構成図で用いているキャラクタ画像（図中の青色の丸で囲まれた領域）を別の画像に置き換えた版のスライドを探せ」と指示した．図／画像だけが異なるスライドが多数存在し，そのタイトルである“高速な不正利用画像検出システム MIRACLE Detector”を検索クエリとしても，13 件の文書，14 枚のスライドがヒットする．テキストの内容は同一であり，これ以上のキーワードによる絞り込みはできない．

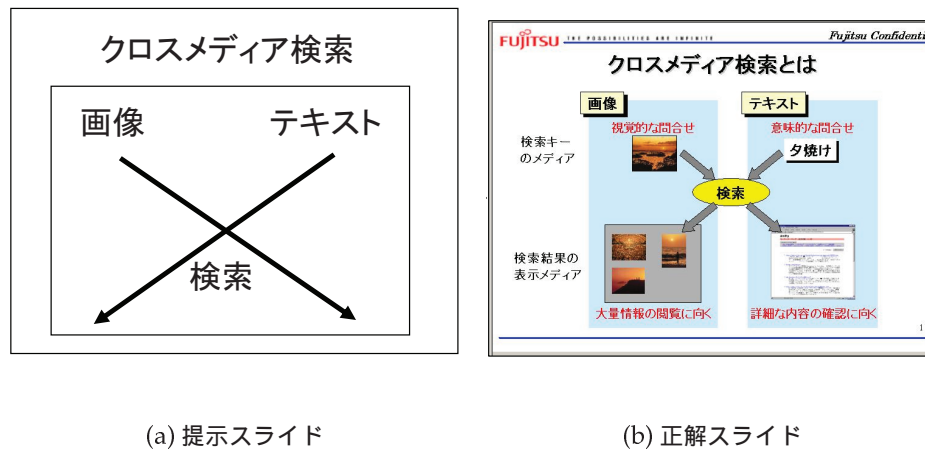


図 2.26 デジタル文書検索実験 (Task3)

Task3 見た目の印象だけを覚えている場面

過去に見たことがあるが、見た目の印象だけを覚えている場面を想定して、被験者に図 2.26(a) を提示し、「見た目の印象だけを覚えているこの図のようなスライドを探せ」と被験者に指示した。図 2.26(b) が実際の検索対象のスライドであるが、被験者に提示したのは、図 2.26(a) だけである。検索クエリとして“クロスメディア検索”、“画像”、“テキスト”の論理積を入力しても、43 件の文書（文書に含まれるスライド総数は 500 枚程度）がヒットし、その中の 110 枚のスライドがこれらのキーワードを同時に含んでいる。

実験結果と考察

図 2.27 に、三つの検索タスクにおける被験者 10 名の平均検索時間を示す。

全文検索システムは、Task1 では、適切なキーワードを入力できさえすれば十分に絞り込めるので、パノラミック検索システムよりも速かった。しかし、キーワードによる絞り込みが難しい Task2, Task3 では、文書を逐一チェックせざるを得ず、2.5 倍以上の検索時間を要している。特に、Task3 では、多数の文書のチェックが必要であり、10 名中 3 名の被験者は検索を断念した（この場合、断念した時点までを検索時間とした）。もし、全文検索システムを利用してスライド単位で検索したとすれば、Task2, Task3 では、文書単位の検索よりも多数のスライドを逐一開いてチェックする必要がある、文書単位の検索よりも多くの検索時間を要するものと思われる。

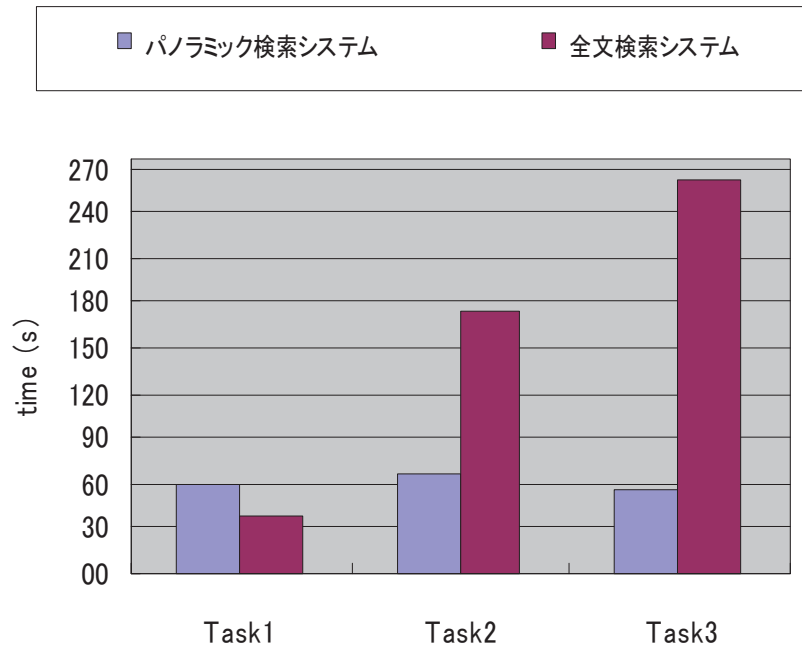


図 2.27 デジタル文書検索実験の結果

一方、パノラミック検索システムでは、Task2, 3 のような場面において、クロスメディア検索および仮想三次元空間への配置機能を用いてスライド全体を概観しながら比較でき、その操作に未習熟な被験者であっても、全文検索システムよりも約 3~5 倍高速に検索できた。

Task1 では全文検索システムに劣ったものの、実験で課した検索タスクは、わずかに二つのキーワードのみで十分に絞り込める極端な場合であり、Task2, 3 での検索時間を考慮すると、パノラミック検索システムの優位性は主張できるであろう。また、被験者によるパノラミック検索システムの操作手順を観察した結果では、スライド全体を仮想三次元空間に配置しウォークスルーする機能だけしか使用していない被験者が多かった。これは、被験者がパノラミック検索システムに習熟していないことやこれらの機能だけでも検索対象を十分に絞り込めたことによるものと考えられる。さらなる検索効率の向上に寄与すると思われる探索機能などについては、実験 3 でその有効性を検証することにする。

本実験での検索タスクは、デジタル文書検索の場面に限らず、“画像と関連テキスト”の対で表現されたマルチメディア・コンテンツの検索場面全般にあてはまるものであり、これらを検索する場合でも、本実験と同様の結果になると考える。

2.8.2 実験 2：映像の検索

実験 1 では、パノラミック検索システムのクロスメディア検索および仮想三次元空間への配置機能の有効性を検証した。ここでは、仮想三次元空間への配置方法の有効性を検証するために、パノラミック検索手法による映像検索システムを用いて、シーン代表画像の一覧表示平面への配置方法の違いによる検索時間の比較、および、早送り再生による検索時間の比較実験を行った。

実験データと実験方法

実験データは、一時間の TV 番組を録画した映像で、天気予報やニュースの番組が続く、比較的多くの CM が流れているものである。この映像データから、上述したように、シーン変化点を自動的に抽出して得られた 646 枚のシーン代表画像を、“HSI 色ヒストグラム特徴に基づく配置”、“時系列配置（シーン代表画像を時間順に折り返して二次元配置したもの）”、“ランダム配置”の三通りの方法で、仮想三次元空間の一覧表示平面上に配置した。

検索タスクは、クローズドキャプションなどの関連テキストは用いずに、シーン代表画像群が配置された空間をウォークスルーするだけで、提示された検索すべきシーン代表画像（標的画像と呼ぶ）を検索することである。実験では、パノラミック検索システムの利用経験の少ない 6 名を被験者として、各配置に対する検索時間を計測した。また、映像の内容を見逃さずに視認できる再生速度であった 3 倍速の早送り再生による検索時間との比較も行った。

図 2.28 に、実験に用いた標的画像を示す。

標的画像として、際立った代表的な色を持つ画像（A, B）と際立った代表的な色を持たない画像（C）を設定した。646 枚のシーン代表画像群の中では、画像 A と色特徴が類似したものは少ないが、画像 B と類似したものは多い。また、実験では、検索過程の観察も行った。

図 2.29 に、シーン代表画像の“HSI 色ヒストグラム特徴に基づく配置”と標的画像 B を示す。

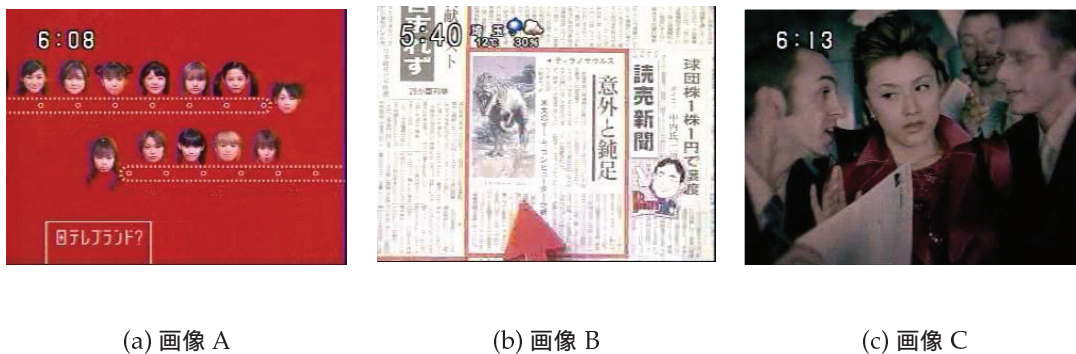


図 2.28 映像検索実験の標的画像

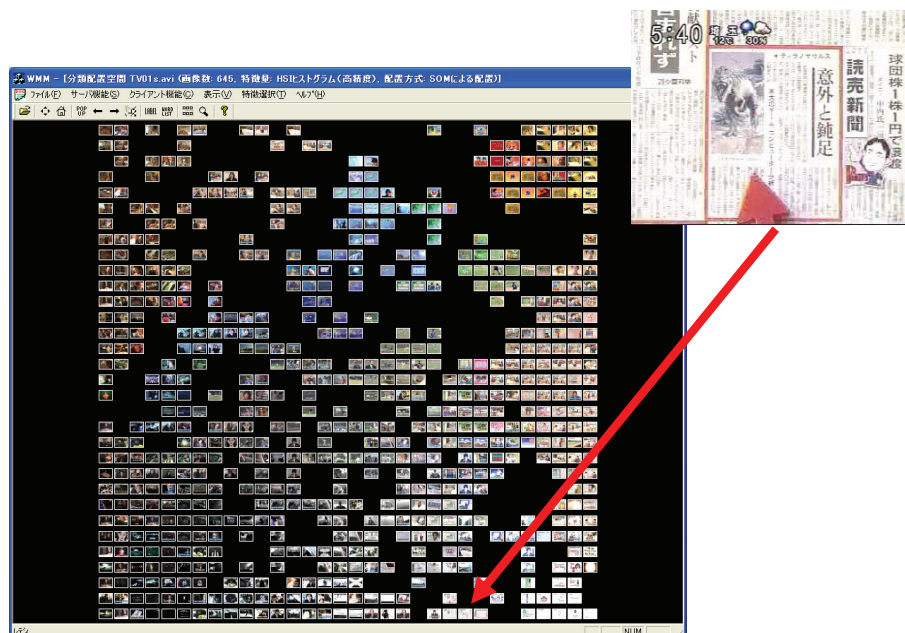


図 2.29 HSI 色ヒストグラム特徴に基づく配置と標的画像 B

実験結果と考察

図 2.30 に、三つの標的画像の各々について、三通りの配置方法の合計 9 通りに対する、被験者 6 名の平均検索時間を示す。また、図 2.31 に、これら 9 通りすべての平均検索時間と 3 倍速の早送り再生における平均検索時間（計算値）との比較を示す。

際立った代表的な色を持つ標的画像 A については、配置方法に拘わらず、すべての被験者が 10 秒以内で検索できた。標的画像 B については、色特徴が類似した画像が多いため、被験者は標的画像が存在すると推測した領域に近づいては詳細をチェックするという

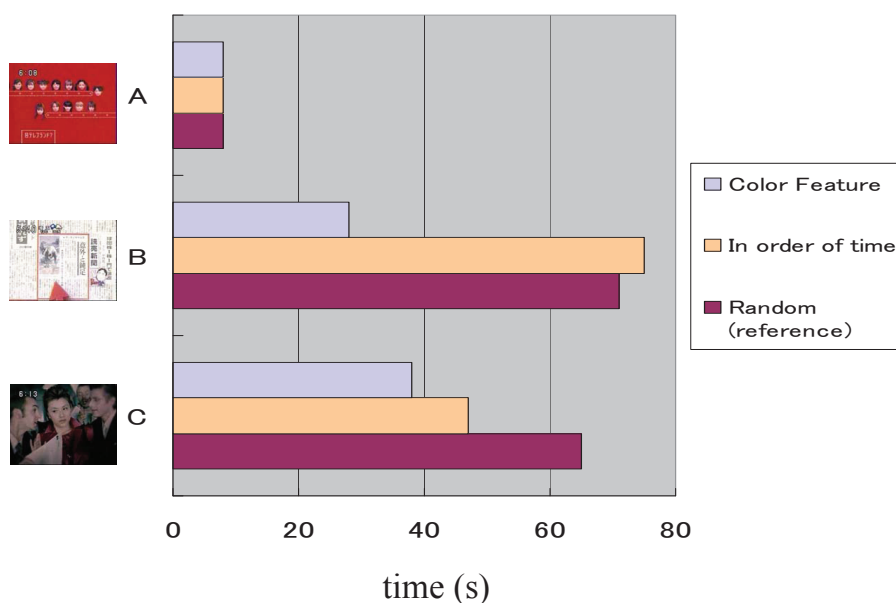


図 2.30 映像検索実験の結果（配置方法による比較）

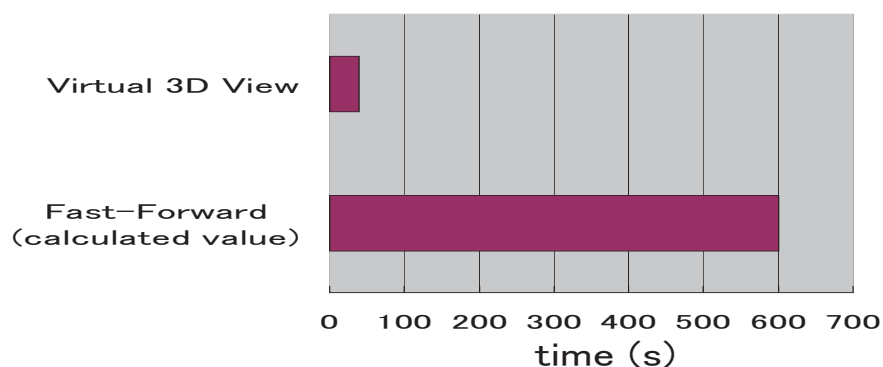


図 2.31 映像検索実験の結果（3 倍速再生との比較）

操作を繰り返すことに大半の時間を費やした．HSI 色ヒストグラム特徴に基づく配置では，標的画像の周辺に類似画像が集まり，そこを集中的に探すことができるので，この配置方法による検索が最も速い．標的画像 C については，HSI 色ヒストグラム特徴に基づく配置が最速ではあるが，標的画像 B ほどの優位性は現れなかった．

本実験の結果，HSI 色ヒストグラム特徴に基づく配置が最も効果的であり，平均して，他の配置の約 2 倍の速さで標的画像を検索可能であることが確認できた．この結果は，画像特徴の類似性に基づく配置が他の配置方法よりも効果的であると一般化しても問題ないであろう．また，3 倍速の早送り再生による検索時間との比較では，HSI 色ヒストグラム特徴に基づく配置は，約 20 倍の速さでの検索を実現している．早送り再生による検索時

間は、標的画像が映像データ内のどのフレームに出現するかによるが、すべてのフレームに出現とした場合、一時間の映像を3倍速の早送り再生で標的画像を検索すると、その平均検索時間は600秒となる。これに対して、三通りの配置方法のうち、最も検索に時間がかかったランダム配置でも、平均検索時間は50秒程度であり、時間軸を持つ映像を仮想三次元空間に展開して検索するパノラミック検索システムの有効性が確認できた。

2.8.3 実験3：画像の検索

実験1では、パノラミック検索手法のクロスメディア検索および仮想三次元空間への配置機能の有効性を検証した。実験2では、仮想三次元空間への配置方法の有効性を検証した。ここでは、パノラミック検索手法の探索機能の一つである絞込み検索機能の有効性を検証するために、実験2と同様に、パノラミック検索手法による映像検索システムを用いて、絞込み検索機能を使用する場合と使用しない場合とで、検索時間の比較実験を行った。

実験データと実験方法

まず、実験データとして使用するために、エジプトの遺産などを紹介したクローズドキャプション付きのTV映像（約一時間の映像）から、シーン変化点を自動的に抽出し、そのシーンに付与されているクローズドキャプションを関連テキストとして持つ190枚のシーン代表画像を得た。これに、2.4節で述べたWeb検索システムで収集したエジプトに関する関連テキスト付きの画像を加え、合計473枚の関連テキスト付きの画像を実験データとした。

検索タスクは、実験2で最も効果的な配置であったHSI色ヒストグラム特徴の類似性に基づく配置により、実験データの画像群を一覧表示平面に配置した状態を初期状態として、提示された標的画像を検索することである。実験では、20名の被験者を絞込み検索機能を使用する被験者10名と、使用しない被験者10名の二組に分け、標的画像の検索時間を計測した。また、検索過程の観察も行った。

図2.32に、被験者に提示した標的画像と関連テキストを示す。また、図2.33(a)に、画像群をHSI色ヒストグラム特徴に基づいて一覧表示平面に配置した初期状態を、図2.33(b)に、Task1の実験において絞込み検索機能を使用した直後の画面を示す。

標的画像として、際立った代表色を持つ画像（Task1）、よく知られているスフィンクスやピラミッドの画像（Task2、Task4）、際立った代表色を持たない画像（Task3）を設定した。Task2とTask4の標的画像は、互いに類似した画像であるが、スフィンクスの画像が2枚だけであるのに対し、ピラミッドの画像は十数枚ある。

広大な砂漠の広がる大地に、豊かな水をもたらす大河ナイル。



(a) Task1

スフィンクスの顔は、カフラー王自身に真似て造られたといわれていますが、真実は謎に包まれています。



(b) Task2

市内南部にそびえる中世の城塞・シタデル



(c) Task3

ジュセル王のピラミッドは、クフ王のピラミッドより、更に80年ほど古い世界最古のピラミッドです。初期のピラミッドは、このように階段状に建設され、玄室は地下に造られていました。

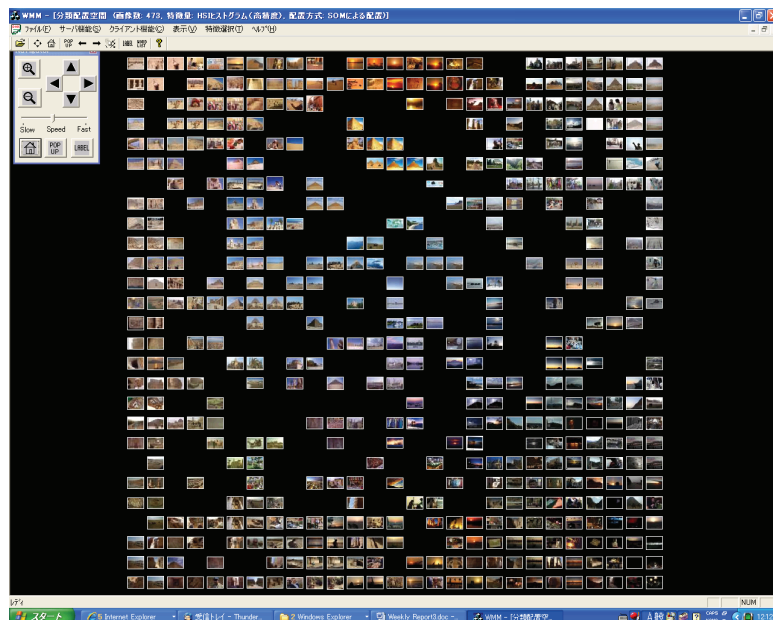


(d) Task4

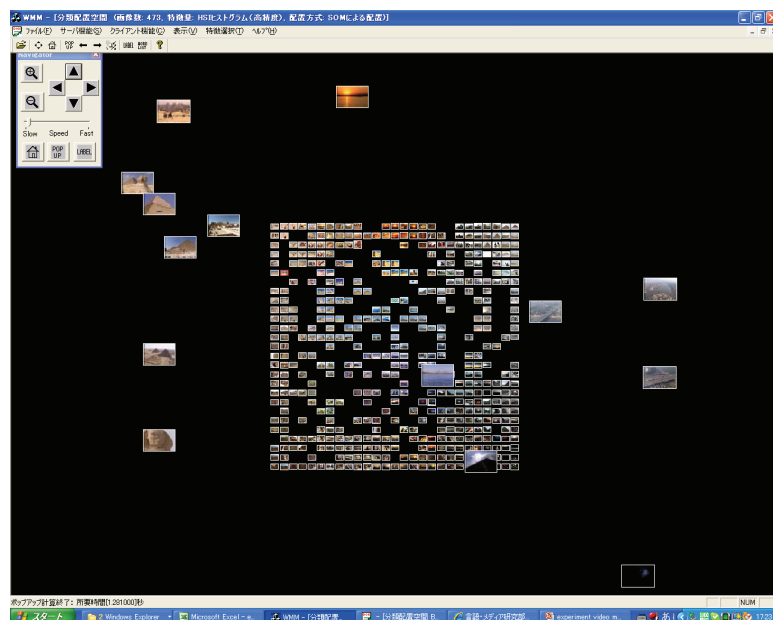
図 2.32 画像検索実験の標的画像

実験結果と考察

図 2.34 に、Task1～Task4 の各々に対する、絞り込み検索機能を使用した被験者の組と使用しなかった被験者の組の平均検索時間を示す。



(a) HSI 色ヒストグラム特徴に基づく配置



(b) 絞り込み検索

図 2.33 画像検索実験 (Task1)

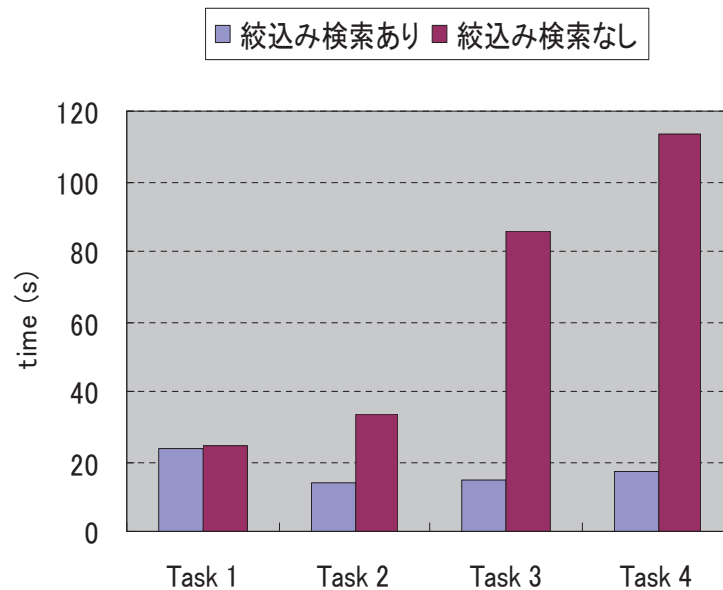


図 2.34 画像検索実験の結果

標的画像が際だった代表的な色を持つ Task1 では、検索時間に差は見られないが、Task2 ~ Task4 については、明らかに、絞込み検索機能を使った方が検索時間は短くなることが分かった。絞込み検索を使用した被験者は、Task1 ~ task4 のいずれにおいても、適切なキーワードを入力しており、ほぼ 1 回の絞込み検索で、標的画像を検索できた。Task1 ~ Task4 の絞込み検索を使用した被験者の組の平均検索時間は、ほぼ 1 回の絞込み検索で標的画像を検索するのに要した時間であり、同程度の検索時間となっている。一方、絞込み検索を使用しなかった被験者の組の平均検索時間は、Task1 ~ Task4 でかなり異なっている。Task2、Task4 の標的画像は、互いに類似しており、一覧表示平面上で近くの位置に配置されているが、Task2 では、スフィンクスの画像は 2 枚しかないため、一旦、スフィンクスの画像を見つけると、すぐに標的画像を検索することができる。ところが、Task4 では、ピラミッド画像を見つけても、多数のピラミッド画像が存在するため、標的画像を特定するのに時間を要している。Task3 については、標的画像は際だった代表的な色を持たない画像であるが、Task4 ほどに、類似した画像が多くないため、Task4 よりは検索時間は短くなっている。

このように、検索時間は、標的画像がどのような性質の画像であるかに依存するが、いずれの場合も、絞込み検索機能を使用した方が検索時間は短くなっており、絞込み検索機能の有効性は確認できたといえる。

2.8.4 実験のまとめ

実験1では、パノラミック検索手法のクロスメディア検索および仮想三次元空間への配置機能の有効性が確認できた。実験2では、仮想三次元空間への配置機能の中でも、画像特徴の類似性に基づく配置の有効性が確認できた。また、時間軸を持つ映像を仮想三次元空間に展開して検索するパノラミック検索手法の有効性も確認できた。実験3では、探索機能の一つである絞込み検索の有効性が確認できた。

関連研究との比較という観点では、パノラミック検索手法は、すべてのインタラクティブな検索過程でクロスメディア検索と仮想三次元空間を利用した視覚化を実現しており、検索対象に応じてさまざまな機能を自由に使い分けながら、より効率的な検索を実現できると考える。

2.9 まとめ

本章では、人間とシステムとの間に存在する Semantic Gap の解消に向けて、人間を積極的に検索ループの中に取り込むことを検討し、1) クロスメディア検索、2) 検索結果および検索過程の仮想三次元空間による視覚化、を特徴とする、インタラクティブなマルチメディア・コンテンツのパノラミック検索モデルとそれに基づくパノラミック検索手法を提案した。

提案手法は、マルチメディア・コンテンツを構成する各メディアの特性および各メディア間の関連性を有効に活用し、コンテンツが画像とテキストで構成される場合には、画像による視覚的な検索とテキストによる意味的な検索とを兼ね備えた情報検索手段を提供する。また、検索結果としてのコンテンツ集合やインタラクティブな検索過程を仮想三次元空間上で人間に分かりやすく視覚化することによって、人間の視覚認識能力を活用した効率的な検索が実現できる。

提案手法は、主に画像とテキストから構成されるマルチメディア・コンテンツを対象として、検索対象を“画像と関連テキスト”の対として表現する。検索対象とするマルチメディア・コンテンツがこの図式で表現できれば、パノラミック検索手法の枠組みがそのまま適用できる。その適用例として、Web 検索、映像検索、画像検索、デジタル文書検索、図面検索、3D-CAD モデル検索システムを示した。

提案手法を適用したデジタル文書検索システム、映像検索システムによる評価実験を行い、パノラミック検索手法の有効性を確認した。

第 3 章

検索手法から解析手法への拡張

3.1 はじめに

本章では，パノラミック検索手法の枠組みを，大量のマルチメディア・コンテンツからの知識発見を支援する解析手法の枠組みへと拡張し，画像群からの知識発見を支援するビジュアル解析手法として，パノラミック解析手法を提案する．

画像群からの知識発見では，無数に定義できる視覚的特徴がある中で，有用な知識に關与する視覚的特徴を発見することが本質的な課題である．

この課題を解決するために，提案するパノラミック解析手法は，人間を解析ループの中に取り込み，人間の高度な視覚認識能力を有効に活用し，人間による知識発見を支援する立場を採ることにより，インタラクティブに有用な知識に關与する視覚的特徴さらには有用な知識を発見しようとするアプローチである．

3.2 関連研究

画像群からの知識発見の関連研究として，脳の f-MRI 画像からの知識発見 [12]．気象画像からの知識発見 [13]，[14]，[15]，医療画像からの知識発見 [16]，航空写真からの知識発見 [17] など，さまざまな応用分野に関する研究がある．例えば，月本ら [12] の脳の f-MRI 画像からの知識発見手法は，脳の f-MRI 画像とさまざまな動作行為との関係から，脳の活動部位と対応する脳の機能を発見する手法を提案している．

これらの画像群からの知識発見の研究では，無数に定義できる視覚的特徴がある中で，有用な知識に關与する視覚的特徴が既知であることを前提としている．例えば，月本らの研究では，さまざまな動作行為と関連するのは，f-MRI 画像を格子上に区切った領域の活

性化パターンの特徴であるということを前提としている．このような前提を置くことができれば，画像群からの知識発見問題は，その既知である視覚的特徴を画像特徴として数値化して抽出することにより，データマイニングの問題に帰着でき，回帰分析やアソシエーション・ルールの抽出 [41] など既存のデータマイニング手法が適用できる．

しかし，現実には，有用な知識に関与する視覚的特徴が既知でない問題も数多く存在しており，また，有用な知識に関与する視覚的特徴が既知であったとしても，解析システムがその視覚的特徴を画像特徴として自動的に抽出できるとは限らない．特に，視覚的特徴がより高次の意味的なレベルのものであれば，その特徴を自動的に抽出することは，一般に困難である．

一方，Langley は，自然科学分野での知識発見における人間と計算機との協調の有用性を主張し，いくつかの成功事例を報告している [42] ．

パノラミック解析手法は，有用な知識に関与する視覚的特徴が既知であることを前提とせず，人間と計算機との協調に基づく解析手法として具現化している点で，関連研究とは異なっている．

3.3 パノラミック解析手法

解析対象とするコンテンツ集合は，パノラミック検索手法の表現図式と同様に，その要素が画像とテキスト／数値データとの対で表現されているものとする．なお，画像と対で表現されるテキスト／数値データを画像からみて“関連データ”と呼び，それがテキストの場合には“関連テキスト”と呼ぶ．

画像群からの知識発見の問題は，画像の視覚的特徴と関連データとの間に成り立つ関係（知識）を発見することと捉えることができる．この問題に対して，高度な視覚認識能力を有する人間が画像群を眺める過程を中心に据え，パノラミック検索手法のさまざまな機能を用いて知識発見を支援しようというのが，本手法の基本的な考え方である．

例えば，バッグの商品画像（画像）や売上データ（関連データ）などのコンテンツを用いた市場調査を想定するならば，パノラミック検索手法の関連データに基づく配置機能を用いると，類似した売上データを持つバッグは，仮想三次元空間の一覧表示平面上の近くに集まって配置される．このとき，近くに集まって配置されたバッグの商品画像に共通する視覚的特徴が存在すれば，その商品画像の視覚的特徴（例えば，「ページュで手提げの長い形状」と売上データ（例えば，「10代の女性に対する売上」と）との間に何らかの関係が成り立っていることになる．したがって，「ページュで手提げの長い形状のバッグは，10代の女性に売れている」といったバッグの商品画像と売上データとの間に成り立つ関係，

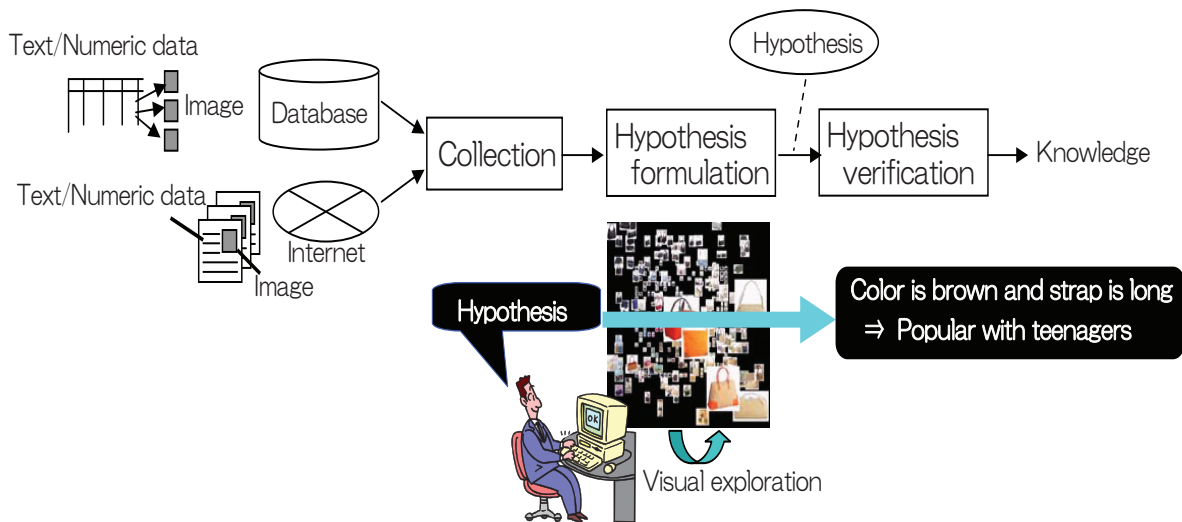


図 3.1 パノラミック解析手法の概要

すなわち，知識を発見できる．

図 3.1 に，コンテンツ収集，仮説設定，仮説検証の三つの過程から成る，パノラミック解析手法の概要を示す．

3.3.1 コンテンツ収集

コンテンツ収集では，適用分野ごとの目的に合ったコンテンツ集合，すなわち，画像と関連データの対で構成されるコンテンツ集合を収集する．市場調査の例では，個々の商品についての時期別，購買者の年齢別，販売地区別などの売上を表す数値データや商品の紹介を記述したテキストが関連データとして，商品の画像と関連付けられて収集される．

3.3.2 仮説設定

仮説設定では，ユーザが画像群を眺めることで，画像の視覚的特徴と関連データとの間に成り立つ関係，すなわち，知識の候補を発見する．ここでは，この知識の候補を仮説と呼ぶことにする．この仮説設定，すなわち，知識候補の発見を支援するために，コンテンツをその関係性に基づいて仮想三次元空間に配置する．ここでは，コンテンツ間の類似関係を仮想三次元空間の位置関係で表現する場合に焦点をあてる．

パノラミック検索手法の配置機能は，コンテンツ間の類似関係を仮想三次元空間上の位置関係に対応付けるものであり，画像および関連データから自動的に抽出した特徴ベクト

ル間の類似関係を仮想三次元空間上の距離に対応付けて配置の位置関係を決定する．例えば，関連データから抽出した特徴ベクトル間の類似度に基づいて計算した仮想三次元空間上の配置結果に従い，その関連データの対として関連付けられた画像を配置して表示する．その結果，類似した画像群が近くに集まって配置されるならば，画像と関連データとの間に何らかの関係が存在することになり，ユーザはその関係を容易に把握して，知識の候補である仮説を設定できる．仮説（知識）の具体的な表現として，例えば，式 (3.1) で表されるアソシエーション・ルール [41] を用いることができる．

$$F \Rightarrow M \quad (3.1)$$

但し， F は，「画像の色がベージュである」といった画像特徴に関する条件を表し， M は「売上が1億円以上である」とかといった関連データに関する条件を表す．

パノラミック解析手法は，2.4.3 節で述べた画像の色特徴や形状特徴による配置やテキスト特徴による配置機能の他にも，仮想三次元空間に画像群を関連データの数値データの昇順に配置する機能や，仮想三次元空間の各軸に関連データの数値データを割り当てて画像群を配置するといった知識発見を支援する機能を有しており，ユーザはインタラクティブに画像群の配置を切り替えながら，さまざまな観点から知識の候補を発見し，仮説を設定することができる．

3.3.3 仮説検証

仮説検証では，設定した仮説を知識として活用するために，統計的な指標などを用いて仮説を検証し，発見した知識候補の中から，有用な知識を抽出する．例えば，仮説（知識）の具体的表現として，アソシエーション・ルールを用いる場合には，ルールの重要度を測る尺度として *support* と *confidence* を算出して評価する．

support と *confidence* は，以下のように定義される．

$$support = N(F \wedge M) / T \quad (3.2)$$

$$confidence = N(F \wedge M) / N(F) \quad (3.3)$$

但し， T は，画像と関連データの対で表現されたコンテンツの総数， $N(F \wedge M)$ は，画像特徴の条件 F と関連データの条件 M の両方を満たすコンテンツの数， $N(F)$ は，画像特徴の条件 F を満たすコンテンツの数である．

3.4 実験

パノラミック解析手法を実装した知識発見支援システム（以下、パノラミック解析システムと呼ぶ）を開発し、その有効性を検証するために、以下の三通りの実験を行った。

3.4.1 実験 1：“陶磁器”画像群からの知識発見

実験 1 では、文化財を用いた発見学習 [43]、[44]、[45] などの教育分野への適用を想定して、陶磁器画像群からの知識発見の実験を行った。

実験データと実験方法

実験データとして、Web から収集した 166 枚の陶磁器の画像と関連テキストの対で構成されるコンテンツ集合を使用した。関連テキストには、「この壺の特徴は... にあり、製作者は... であり、1600 年代後半に作られた。また、...」といった自由文で、各陶磁器の製作年代、製作者、製作技法などの属性が記述されている。テキスト特徴として、単語頻度特徴ベクトルを使用した。この要素は、関連テキスト中に生起する単語の中で助詞等の不要語を除いた単語の tf・idf 法 [26] に基づく重要度である。実験では、陶磁器に関する専門的な知識を持たない被験者 1 名が、パノラミック解析システムを用いて知識発見を試みることとし、その過程を観察した。

実験結果と考察

被験者は、関連テキストの単語頻度特徴ベクトル間の類似度に基づいて陶磁器の画像群を配置した。図 3.2 に、その配置結果を示す。この結果を眺めると、左下の楕円で囲まれた領域に「白と青の絵や模様」の類似した視覚的特徴を持つ画像群が、また、右上の楕円の領域に「円形で黄金色」の類似した視覚的特徴を持つ画像群が集まっている。したがって、これらの陶磁器の視覚的特徴とその配置に用いた関連テキスト特徴との間に何らかの関係が存在することは分かるが、これらの領域内の関連テキスト集合の共通的な特徴が何であるかは分からない。仮説としてアソシエーション・ルールを設定するためには、関連テキスト集合に共通的なテキスト特徴を知る必要があり、システムには、特定の関連テキスト集合の中で重要となる単語を抽出する“重要単語抽出機能”が実装されている。ユーザが表示画面上で関心のあるコンテンツ集合を選択すると、それらの関連テキストに現れる各単語の重要度が計算され、重要度順にリスト表示される。

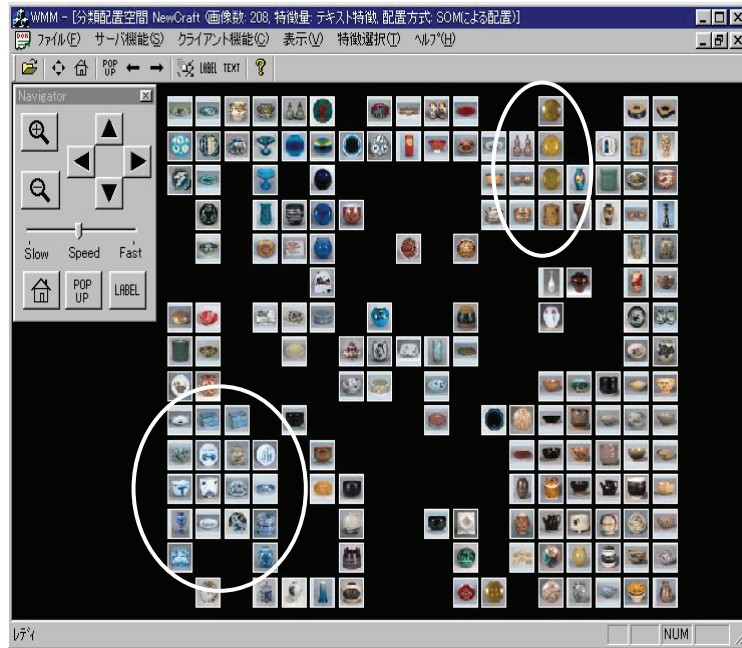


図 3.2 テキスト特徴による陶磁器画像の配置

被験者は、「白と青の絵や模様」という視覚的特徴を持つコンテンツ集合を GUI を用いて囲むことにより選択し、重要単語抽出機能を実行して、図 3.3 に示すように、重要単語のリストを表示した。このリストから、選択したコンテンツ集合の関連テキストに共通的な特徴として、“染付”という単語が最も重要であることを知り、以下の仮説 $H1$ を設定した。

$$H1: F_{white \& blue} \Rightarrow M_{sometuke} \quad (3.4)$$

但し、 $F_{white \& blue}$ は、陶磁器が「白と青の絵や模様」を持つことを、 $M_{sometuke}$ は、関連テキスト中の重要単語として“染付”を持つことを表す。すなわち、仮説 $H1$ は、「ある陶磁器の画像が白と青の絵や模様を持つならば、その陶磁器は“染付”と強い関連がある」というルールを表している。

また、被験者は、パノラミック解析システムの絞込み検索機能を用いて、図 3.4 に示すように、“染付”という重要単語を持つ画像を手前に大きく表示し、この仮説を確認した。このルールを検証した結果、 $support$ は 0.07、 $confidence$ は 0.82 であった。なお、“染付”とは陶磁器の製作技法の一つである。

本実験結果から、パノラミック解析システムを用いて、陶磁器の専門的知識を持たない被験者でも、高い $confidence$ の値を持つアソシエーション・ルールを発見できることが確認できた。

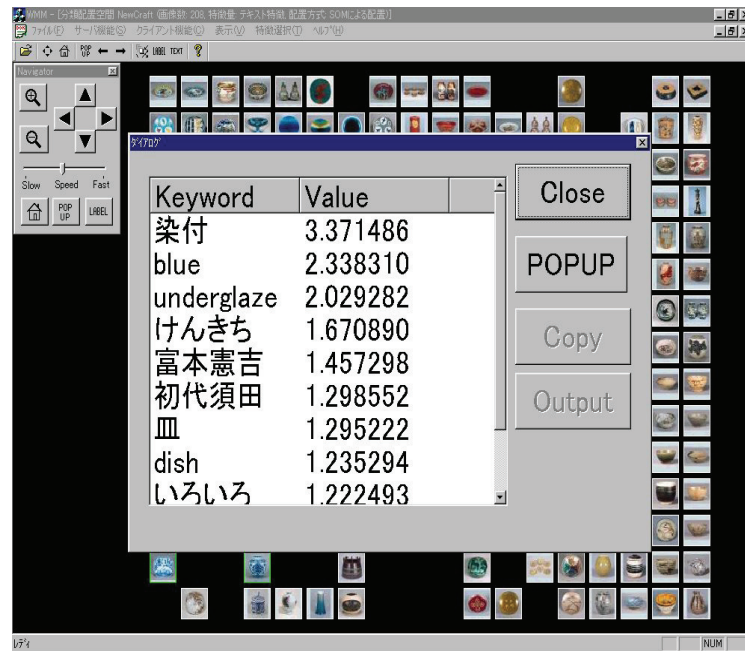


図 3.3 重要単語のリスト表示

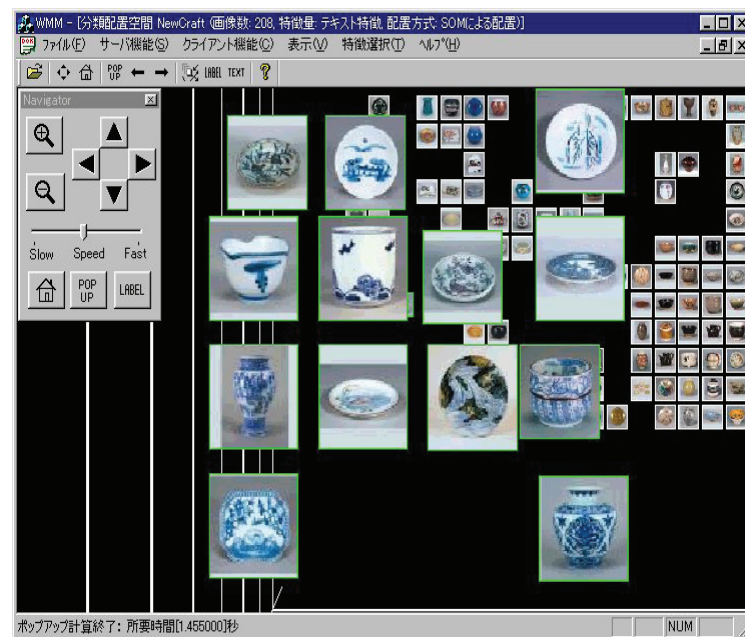


図 3.4 “染付”による絞込み検索結果

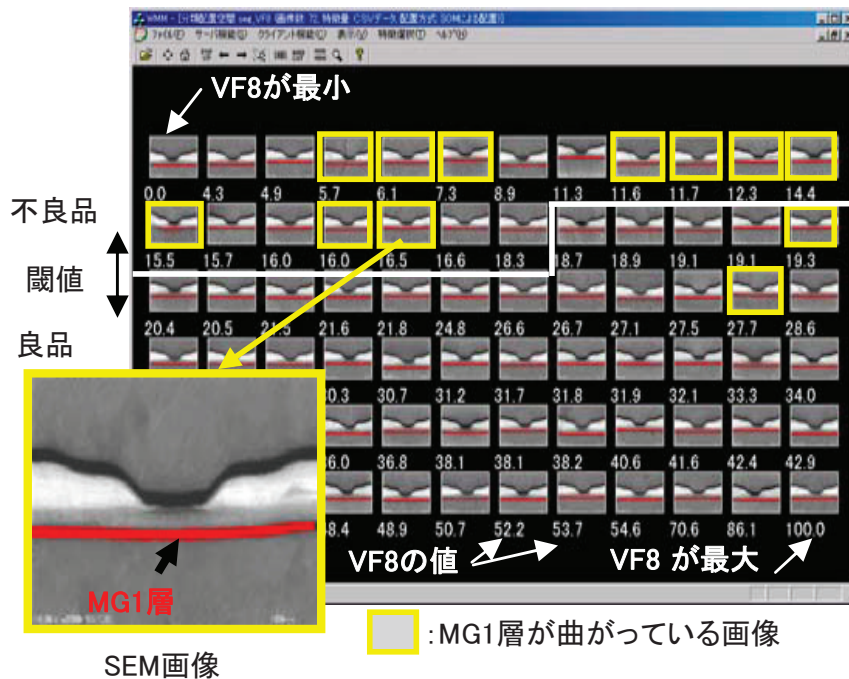


図 3.5 関連データ“VF8”による配置

3.4.2 実験2：GMR ヘッド形状の解析

実験2では、設計・製造分野への適用可能性を検討するために、実データを用いた HDD (Hard Disk Drive, HDD) の GMR (Giant Magneto Resistive, GMR) ヘッド形状の分析実験を行った。

実験データと実験方法

GMR ヘッドは、ディスクからデータを読み込むための“リード素子”と、ディスクへデータを書き込むための“ライト素子”から成る。

図 3.5 に、パノラミック解析システムによる解析の画面を示す。図 3.5 の各画像は、SEM (Scanning Electron Microscope, SEM) と呼ばれる電子顕微鏡 (20 万倍) で撮影されたもので、リード素子の ABS (Air-Bearing Surface, ABS) 浮上面 (ディスクと数十 nm の間隔で接する面) を示す。GMR ヘッドは多数の薄膜で構成され、隣接する層の接地面が SEM 画像に輪郭として現れる。この輪郭形状の違いが性能に影響していると考えられている。

ここでの解析の目的は、SEM 画像の輪郭の形状とヘッドの性能との関係を発見するこ

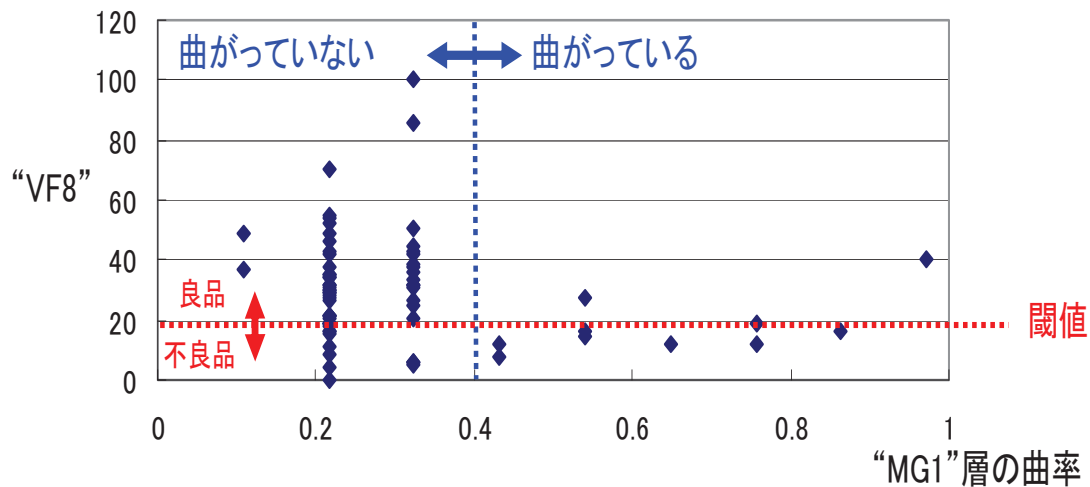


図 3.6 “MG1” 層の曲率と “VF8” との関係

とである．ヘッドの性能は，関連データ “VF8” と呼ばれる出力電圧で表され，“VF8” の値が大きいほど性能が良い．実験では，SEM 画像とヘッドの性能値の対で表現される 72 組のコンテンツ集合を実験データとして用いた．また，この分野の専門知識を持たない被験者 1 名が知識発見を試みることで，発見した知識をヘッド設計の専門家に提示し評価してもらうという方法を使った．

実験結果と考察

まず，被験者は，仮説を設定するために，図 3.5 のように，関連データ “VF8” の昇順に SEM 画像を配置した．図 3.5 の白線は，良品，不良品の閾値（“VF8” = 18.5）を表し，“VF8” が閾値未満であれば，ヘッドは不良品とみなされる．続いて，被験者は，各画像の “MG1” 層の形状に着目し，以下の仮説 $H2$ を設定することができた．

$$H2 : F_{MG1_bend} \Rightarrow M_{defective} \quad (3.5)$$

但し， F_{MG1_bend} は “MG1” 層が曲がっていることを， $M_{defective}$ は “VF8” が閾値未満の不良品であることを表し，仮説 $H2$ は，「“MG1” 層が曲がっているならば，不良品である」というルールを表している．

なお，図 3.5 では，“MG1” 層の形状を目立たせるために，各 SEM 画像の “MG1” 層を赤色で表示している．

図 3.6 は，仮説 $H2$ を検証するために，各 SEM 画像の “MG1” 層の曲率を “MG1” 層の中心線を用いて近似的に求め，その曲率と “VF8” との関係を示したものである．“MG1” 層の曲率が 0.4 よりも大きい場合を曲がっているとみなし，*support* と *confidence* を求め

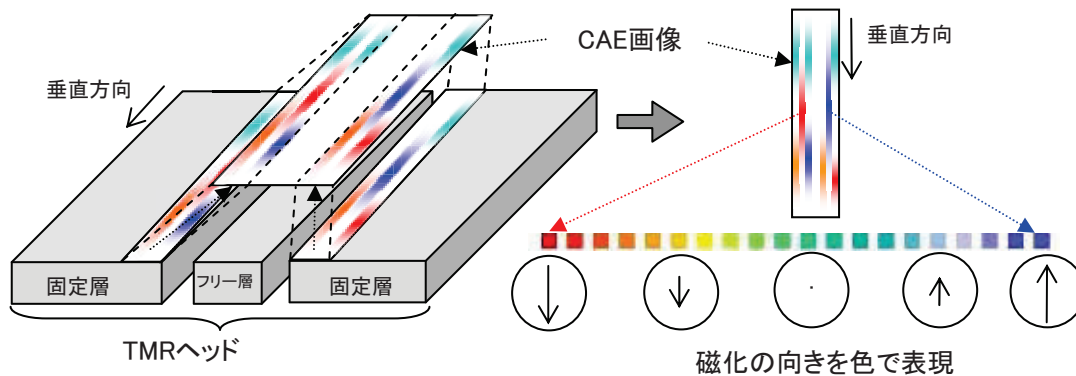


図 3.7 TMR ヘッドとその CAE 画像

ると、それぞれ 0.14 と 0.83 であった。

本実験から、パノラミック解析システムを用いて、GMR ヘッドに関する専門知識を持っていないとも、高い *confidence* の値を持つ仮説 $H2$ を発見できることが確認できた。また、発見した知識に関するヘッド設計者との議論では、発見した知識はこれまでヘッド設計者にも知られていなかった未知の知識であり、新しい知見が得られたことが評価された。

3.4.3 実験 3：TMR ヘッドの磁化ベクトル分布の解析

実験 3 では、設計・製造分野へのもう一つの適用実験として、CAE (Computer Aided Engineering, CAE) による計算機シミュレーション結果を基に、HDD の TMR (Tunneling Magneto Resistive, TMR) ヘッドの磁化ベクトル分布の解析実験を行った。

実験データと実験方法

TMR ヘッドは、超高密度 HDD の読み出し磁気ヘッドとして有望視されている。ここでの解析の目的は、ヘッドの磁化ベクトル分布（磁化の向き）とヘッドの性能との関係を発見することである、図 3.7 に示すように、CAE による計算機シミュレーションの結果得られるヘッド固定層（ハード膜）の垂直方向の磁化ベクトルの分布を、CAE 画像の色を用いて表す。例えば、赤色の領域は、下向きの磁化を、青色の領域は、上向きの磁化を表している。

実験では、ヘッドの CAE 画像とヘッドの性能値の対で表現される 50 組のコンテンツ集合を実験データとして用い、ヘッドの性能の指標として、出力電圧波形の“非対称度”（図 3.8）を用いた。“非対称度”が小さいほど性能が良い。実験 2 と同様に、専門知識を

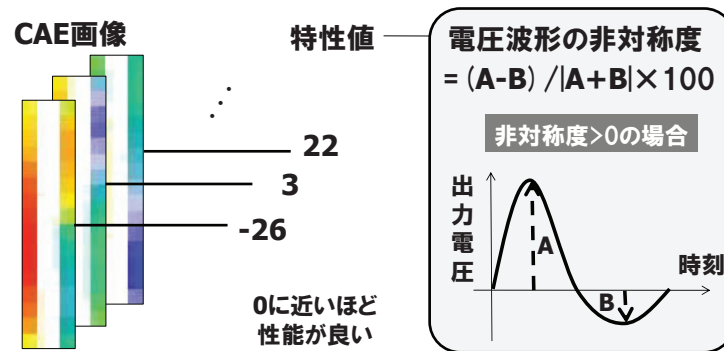


図 3.8 CAE 画像と電圧波形の“非対称度”

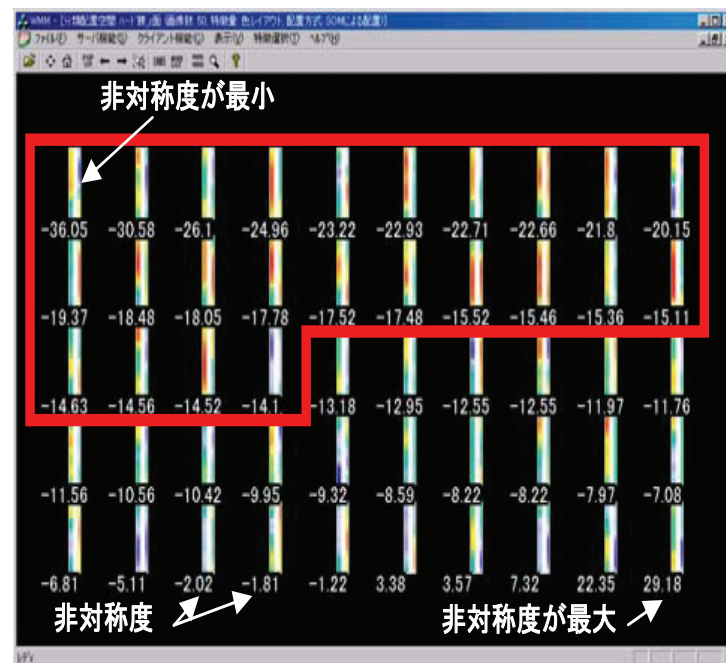


図 3.9 関連データ“非対称度”による配置

持たない被験者 1 名が知識発見を試みることとし、発見した知識をヘッド設計の専門家に提示し評価してもらった。

実験結果と考察

まず、仮説を設定するために、被験者は、図 3.9 に示すように、関連データ“非対称度”の昇順に CAE 画像を配置し、図 3.9 の赤色の枠で囲まれた CAE 画像に注目することで、以下の仮説 H3 を設定することができた。

$$H3 : F_{red} \Rightarrow M_{asymmetry} < -14 \quad (3.6)$$

但し, F_{red} は, CAE 画像が赤色の領域を含むことを表し, $M_{asymmetry < -14}$ は, “非対称度” が “-14” よりも小さいことを表す.

仮説 $H3$ を検証するために, 条件 F_{red} を目視で確認し, $support$ と $confidence$ を求めた結果, $support$ は 0.36, $confidence$ は 0.82 であった.

次に, 被験者は, CAE 画像の色に関する別の仮説を設定するために, 画像の色特徴 (HSI 色レイアウト特徴) による配置に変更した. 図 3.10(a) に, その結果を示す. HSI 色レイアウト特徴が類似した画像が近くに配置されていることが確認できる. 続いて, 被験者は, 関連データである性能値との関係を調べるために, 絞り込み検索機能を使って, 奥行き (Z 軸) 方向に “非対称度” を割り当て, “非対称度” が大きい CAE 画像ほど手前に大きく表示されるように配置した. 図 3.10(b) に, その結果を示す.

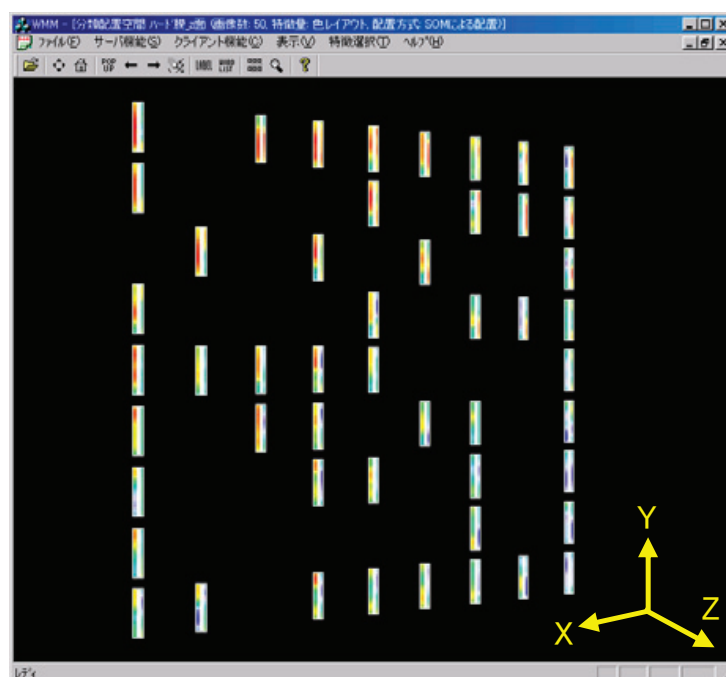
ここで, 被験者は, この配置結果の赤色の枠で囲まれた “非対称度” の大きい CAE 画像に着目して, これらの CAE 画像の右下の領域が青色であることを発見し, 以下の仮説 $H4$ を設定した.

$$H4: F_{blue_in_br} \Rightarrow M_{asymmetry > 0} \quad (3.7)$$

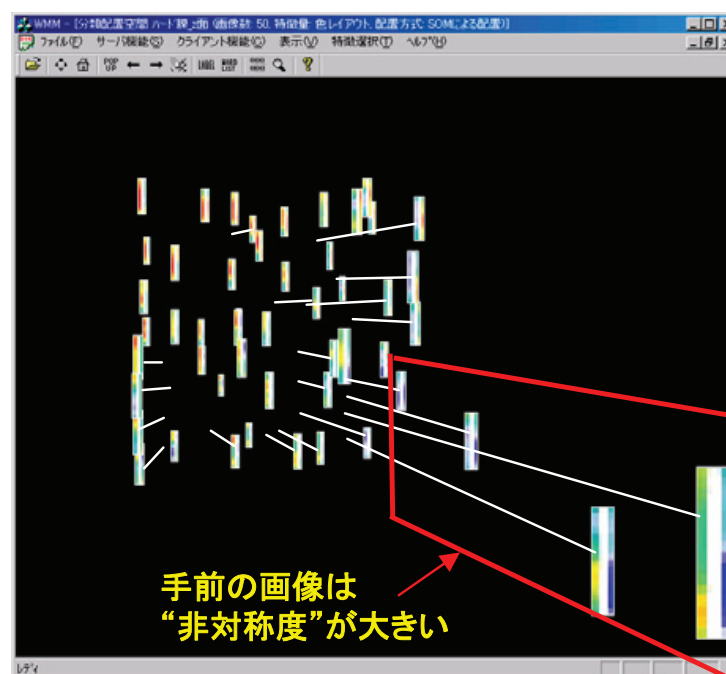
但し, $F_{blue_in_br}$ は, CAE 画像の右下の領域が青色であることを, $M_{asymmetry > 0}$ は, “非対称度” の値が 0 よりも大きいことを表す.

仮説 $H4$ を検証するために, 条件 $F_{blue_in_br}$ を目視で確認し, $support$ と $confidence$ を求めた結果, $support$ は 0.08, $confidence$ は 0.80 であり, 高い $confidence$ の値が得られたことを確認した.

本実験の結果, パノラミック解析システムを用いて, TMR ヘッドに関する専門知識を持っていないとも, 高い $confidence$ の値を持つ仮説 $H3$ および $H4$ を発見できることが確認できた. 発見した知識に関するヘッド設計者との議論では, 仮説 $H3$ については, 既に有用性が確認されている既知の知識であり, 仮説 $H4$ については, ヘッド設計の専門家にとっても未知の知識で, 設計において有用であるとの評価を得た.



(a) 色特徴による配置



(b) 絞込み検索による配置

図 3.10 CAE 画像の色特徴および絞込み検索による配置

3.4.4 実験のまとめ

発見学習などの教育分野を想定した評価実験，設計・製造分野の実データを用いた評価実験のいずれにおいても，その分野の専門家でなくとも，高い *confidence* の値を持つ有用と思われる知識の発見が可能であることを確認でき，この分野への提案手法の適用可能性を示すことができた．特に，設計・製造分野では，画像と関連データとの間に何らかの関係はあるが，具体的にどのような視覚的特徴がどの関連データに影響を及ぼしているかが未知である問題が数多く存在し，有用な知識に關与する特定の視覚的特徴を事前に仮定しないパノラミック解析手法は有効であろう．

3.5 まとめ

本章では，パノラミック検索手法の枠組みを，大量のマルチメディア・コンテンツからの知識発見を支援するための解析手法の枠組みへと拡張し，画像群からの知識発見を支援するビジュアル解析手法として，パノラミック解析手法を提案した．

パノラミック解析手法は，人間を解析ループの中に取り込み，人間の高度な視覚認識能力を有効に活用し，人間による知識発見を支援する立場を採ることにより，インタラクティブに有用な知識に關与する視覚的特徴さらには有用な知識を発見できる．

パノラミック解析手法を実装したパノラミック解析システムを開発し，発見学習などの教育分野を想定した評価実験，および，設計・製造分野の実データを用いた評価実験により，その分野に關する専門知識を持ってなくとも，有用な知識を発見できることを検証し，提案手法の有効性を確認した．

第4章

適合フィードバックの導入

4.1 はじめに

画像検索の分野では，低次の画像特徴と高次の意味概念との間の Semantic Gap を解消するために，適合フィードバックの研究が幅広く進められている．適合フィードバックは，ユーザとシステムとのインタラクションに注目し，ユーザからの検索結果に対する適合／非適合の評価をシステムにフィードバックすることによって，ユーザの検索意図を反映した検索を実現しようとするものである．

一般に，適合フィードバックを実現する上での課題／制約条件として，以下の三つが挙げられる．

- ユーザが適合／非適合と評価できる事例数は限られており，一般に，訓練事例は少ない．特に，正例は少ない．
- 低次の画像特徴は高次元であり，次元数は訓練事例数よりも一般に大きい．
- 実応用で許容できる高速な学習が要求される．

これらの課題／制約条件に対処するために，高次元の特徴に対応でき，少ない訓練事例からの高速な学習に適した SVM を採用し，この SVM の能動学習に基づく適合フィードバック手法が提案されている [9]，[11]．

SVM の能動学習に基づく適合フィードバック手法の実現にあたっては，ユーザに適合／非適合の評価，すなわち，正例／負例のラベリングを要求する事例を能動的に選択するための評価基準の設計が，その手法の性能を左右する重要な鍵となる．

本章では，この能動的な事例選択のための新たな評価基準を設計し，その評価基準を導入した SVM の能動学習に基づく適合フィードバックの新たな手法を提案する．また，提

案手法を CBIR に適用することにより，その有効性を示す．

4.2 関連研究

少数に限られた訓練事例しか獲得できない適合フィードバックの状況では，システムがその制約を受容した上で，いかに訓練事例として有効な事例を選択して獲得するかということが，システムの性能向上，ユーザの負荷軽減に繋がる極めて重要な課題である．

一般に，適合フィードバックにおける少数の訓練事例の問題を解決するために，以下の二つの戦略が採られている．

- 能動学習
- ラベルなし事例の活用

能動学習とは，分類器が最大の情報ゲインを得るために，ユーザにラベリングを要求する事例を能動的に選択するための戦略である．この能動学習の枠組みの下で，SVM を分類器として用いた能動学習に基づく適合フィードバックの研究が進められており，いくつかの手法が提案されている．その代表的なものとして，Tong らの SVM_{Active} (SVM Active Learner) 手法 [9] や Brinker の Angle-Diversity 手法 [11] がある．

Tong らは，特徴空間において，SVM の識別境界近傍の事例を選択することが，最大の情報ゲインを得るための良い選択であることを示し，この事例選択を実現する評価基準として Informative 基準を提案している．

しかし，複数の事例群を選択する場合には，SVM の識別境界近傍の複数の事例を排他的に選択しても，必ずしも最大の情報ゲインが得られるとは限らない．

この問題に対処するために，Brinker の Angle-Diversity 手法は， SVM_{Active} の Informative 基準を改良し，Informative 基準に“選択される複数の事例の多様性”を表す Angle-Diversity 基準を導入した新たな評価基準を提案している．

一方，ラベルなし事例からの学習の枠組みとして，半教師あり学習の代表的な手法の一つである Co-Training の枠組みがある．Co-Training は，二つの異なる観点からの分類器，例えば，SVM と重み付きユークリッド距離モデル (Weighted Euclidean Distance Model) を用いて，ラベルなし事例にラベリングを行い，ラベルあり事例を増やすことで，分類器の分類性能を高めるものである [46]．

4.3 Representative 基準を持つ SVM の能動学習

本論文では，以下の二つの戦略によって， SVM_{Active} の性能を改善する．

- Representative 基準の導入による能動的事例選択の評価基準の強化
- Co-Training の枠組みを利用したラベルなし事例の活用

4.3.1 SVM_{Active} - Informative 基準

Tong らの SVM_{Active} の Informative 基準は，バージョン空間 (Version Space) の大きさを最小にする最も Informative な事例を選択するものであり，そのような事例は SVM の識別境界近傍の点で近似できることが証明されている [9]．

空でない入力空間を χ ，訓練事例の集合を $\{(x_1, y_1) \cdots (x_n, y_n)\} \subset (\chi \times \{-1, +1\})^n$ とすると，Mercer カーネル $k(x, x') = \langle \phi(x) | \phi(x') \rangle$ は，特徴空間 F における内積に対応する入力空間 χ から特徴空間 F への写像 ϕ を定義する．これらの訓練事例が特徴空間において線形分離可能であると仮定すると，バージョン空間は，特徴空間 F における訓練事例を分離する超平面の集合として定義される．

$$V = \{w \in W | y_i \langle w | \phi(x_i) \rangle > 0 \text{ for } i = 1 \cdots n \text{ and } \|w\| = 1\} \quad (4.1)$$

但し， W はパラメータ空間を表す．

特徴空間 F とパラメータ空間 W との間には双対性があり，特徴空間 F の点はパラメータ空間 W の超平面に対応し，パラメータ空間 W の点は特徴空間 F の超平面に対応する．定義により，パラメータ空間 W の点は特徴空間 F の超平面に対応することは明らかである．また，訓練事例 (x_i, y_i) を分離する超平面は， $y_i \langle w | \phi(x_i) \rangle = \langle w, y_i \phi(x_i) \rangle > 0$ を満たさなければならない． $y_i \phi(x_i)$ はパラメータ空間 W の超平面の法線ベクトルであり，この点はパラメータ空間 W の超平面 $w \cdot y_i \phi(x_i) = 0$ を定義し， $w \cdot y_i \phi(x_i) > 0$ はパラメータ空間 W の半空間を定義する．

上記の定義から，学習とは，訓練事例 (x_i, y_i) と矛盾しないパラメータの集合であるバージョン空間 V の探索問題とみなすことができる．この探索を実行するために，Tong らは，バージョン空間 V をできるだけ早く小さくするように，バージョン空間 V を二等分する点を選択する手法を提案した [9]．この選択は，ラベルなし事例 x に対応するパラメータ空間 W における超平面が，バージョン空間 V の中心にどの程度近接しているかを

調べることによって実現できる．パラメータ空間 W における超平面がバージョン空間 V の中心に近いほど，その超平面はバージョン空間 V を二等分する． SVM_{Active} の超平面 w は，バージョン空間 V の中心を近似するので，Tong らの SVM_{Active} の評価基準は，現在の識別超平面に最も近い新しいラベルなし事例を選択することに相当する．

4.3.2 Representative 基準

Tong らの SVM_{Active} の Informative 基準は，最も Informative な事例を選択するための効率的な解法を提供するが，複数の事例群を選択する場合には，問題が生じる．この場合，識別超平面までの距離に基づいて排他的に複数の事例を選択しても，各々の事例がバージョン空間を近似的に二等分するという理論的な根拠は成り立たない．したがって，複数の事例を選択することが，単純に一つ的事例を選択する場合よりも，必ずしもバージョン空間を著しく小さくできるとは限らない [11]．Tong らの SVM_{Active} は，選択される複数の事例間の冗長性を考慮していないため，複数の事例が互いに類似しており，これらの事例により支持される特徴空間は，一般に極めて狭い．したがって，これらの事例で訓練した分類器は，通常，バイアスがかかった不安定なものとなる．

この問題を解決するためのアプローチとして，Brinker は， SVM_{Active} の Informative 基準に“選択される複数の事例の多様性”を表す Angle-Diversity 基準を導入した Angle-Diversity 手法を提案している [11]．Angle-Diversity 基準とは，識別超平面近傍の事例群に対応するパラメータ空間の超平面群間の成す角度に基づいて，多様性のある複数の事例群を選択しようとするものである．

本論文では，さらに広い特徴空間を支持する複数の事例群を選択可能な SVM_{Active} の評価基準として，Representative 基準を提案する．

図 4.1 に，Representative 基準の基本的な考え方を示す．

基本的な考え方は，各回の適合フィードバックにおいて，現在の SVM の識別境界近傍にあるラベルなし事例とそれまでにラベリングされたすべてのラベルあり事例に，動的なクラスタリングを適用し，形成されたクラスタの各々から一つ的事例を選択することによって，複数の事例選択を実現することである．

したがって，Representative 基準により選択される複数の事例は，小さな特徴空間の Representative であり，Informative な性質は保ったまま，選択される事例の多様性を増し，分類器はより大きな情報ゲインを得ることができる．

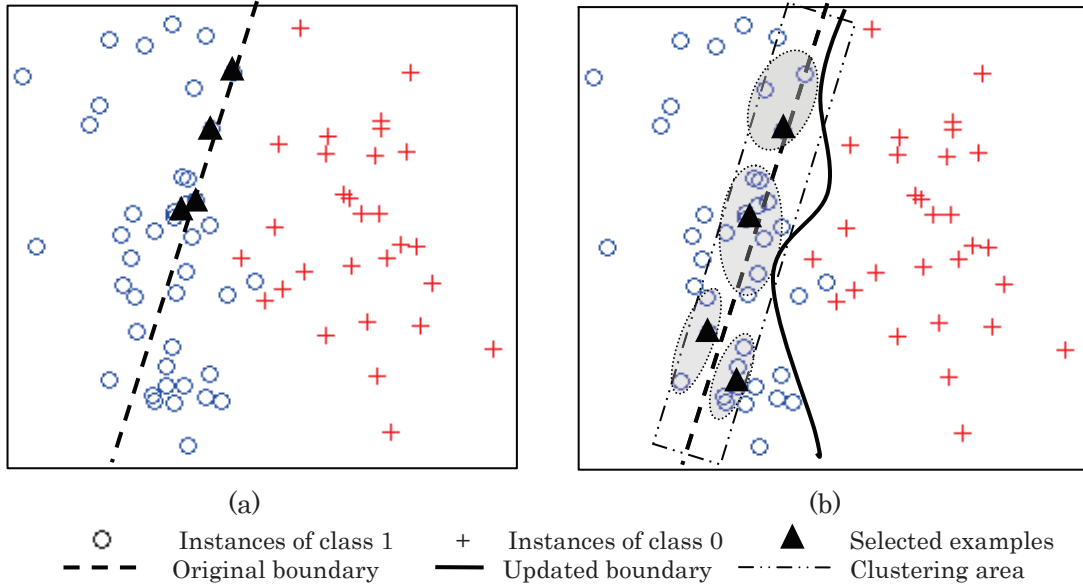


図 4.1 Representative 基準

クラスタリング

本手法では、動的なクラスタリングを実現するために、Normalized Cuts 法 [47] を採用する。Normalized Cuts 法では、クラスタリングはグラフ分割問題として記述される。

クラスタリング対象の N 個の事例が与えられると、重み付き無向グラフ $G = (V, E)$ が構築される。ここで、ノードは事例を表し、各ノード間にエッジが形成される。各エッジの重み w_{ij} は、ノード i と j との間の類似度関数であり、以下のように定義できる。

$$w_{ij} = e^{-d(i,j)/\sigma^2} \quad (4.2)$$

但し、 $d(i, j)$ は、事例 i と事例 j との間の距離、 σ は、スケーリングパラメータである。

一般的に言えば、グラフ分割手法は、グループ内の類似度が高く、グループ間の類似度が低くなるように、ノードをグループ化しようと試みる。グラフ G が二つの互いに素な集合 $A, B, A \cup B = V, A \cap B = \emptyset$ に分割されると仮定すると、その二つの集合を結合するエッジの重みの総和は、この分割を定量化する良い評価尺度であり、 cut と呼ばれる。

$$cut(A, B) = \sum_{u \in A, v \in B} w_{uv} \quad (4.3)$$

$Ncut$ の評価尺度は、次式で定義される [47]。

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \quad (4.4)$$

但し, $assoc(A, V) = \sum_{u \in A, t \in V} W_{ut}$ は, A のノードとそのグラフのすべてのノードとの間のエッジの重みの総和であり, $assoc(B, V)$ も同様に定義される.

Shi and Malik[47] は, $Ncut$ の最小値による分割を見出す問題は, 以下の固有値問題を解き, 二番目に小さい固有値に対応する固有ベクトルを求めることにより実現できることを示した.

$$(D - W)y = \lambda Dy \quad (4.5)$$

但し, W はその要素として $W(i, j) = w_{ij}$ を持つ $N \times N$ の対称行列, D は $d(i) = \sum_j w_{ij}$ を対角要素として持つ $N \times N$ の対角行列である.

実装の問題

Normalized Cuts 法によるクラスタリングを実装する上で, 以下の点を考慮する必要がある.

- どのサブグラフを分割すべきか?
- クラスタの数はどのように決定すべきか? どの時点でクラスタリングを終了すべきか?

一点目について, 本手法では, 最大のノードを持つサブグラフを再帰的に分割する方法を採用する.

二点目について, 最も単純な方法は, クラスタの数を, ユーザにラベリングのために提示する事例の数に設定することである. しかし, この方法は, 事例間の冗長性を生み出す可能性が高い. 通常, いくつかのクラスタは, それ以前の回のフィードバックで既にラベリングされた事例を含むので, これらのクラスタから選択される事例はそのラベリングされた事例と類似したものかもしれない. この場合, これらの新しい訓練事例からはわずかな情報しか得ることができない. また, 別の観点から, クラスタリングや事例の選択の過程において, ラベルあり事例, ラベルなし事例を区別して取り扱わないとすれば, 事例間の冗長性が生み出される可能性が高い.

この問題を回避するために, 本手法では, 以下の評価基準でクラスタリングの過程を制御する.

- ラベルあり事例を含まないクラスタの数が，ユーザにラベリングのために提示する事例の数に等しくなるまで，クラスタリングを繰り返す．

本手法では，ラベルあり事例はそれが属するクラスタの Representative な事例であるということを仮定する．したがって，クラスタリングの過程では，ラベルあり事例を含まないクラスタから事例を選択し，それらの事例をユーザにラベリングのために提示するだけでよい．

クラスタリングの過程は，ラベルあり事例で誘導されるという点で，教師あり学習としてみることもできる．上述したように，クラスタリングを導入する本来の目的は，選択される事例によってカバーされる特徴空間を拡げることであり，この目的は，特徴空間を部分空間に分割し，各々の部分空間から一つの事例を選択することにより達成できる．Representative 基準では，ラベルあり事例を含む部分空間から事例を選択することは許されないで，選択される事例の多様性が保たれる．

要約すると，各回のフィードバックで実行されるクラスタリングの手順は，以下の通りである．

$I = I_N + I_L + I_Q$ をクラスタリング対象の事例とする．但し， I_N は，現在の識別境界近傍の N 個の事例， I_L は，それ以前のフィードバックでラベリングされたラベルあり事例， I_Q は，クエリである．また， T を，ユーザにラベリングのために提示される事例の数とする．

- Step 1. 初期のクラスタ $c_0 = I$ ， $t = 0$ と設定．
- Step 2. 最大のサイズのクラスタ c_i を選択し，Normalized Cuts 法により二つのクラスタ c_m ， c_n に分割する．
- Step 3. 現在のクラスタリストから c_i を除き， c_m ， c_n を追加する．
- Step 4. 以下の二つの条件のいずれかが満たされるならば， t を $t + 1$ に設定する．
 - (a) c_i はラベルあり事例を含まない．
 - (b) c_i はラベルあり事例を含むが， c_m ， c_n のいずれかはラベルあり事例を含まない．
- Step 5. t が T よりも小さいならば，Step 2. へ，そうでなければ，クラスタリングを終了．

4.3.3 結合評価基準

提案手法では、選択される事例が Informative で、かつ、Representative でなければならないという二つの要請を充たすために、Informative 基準と Representative 基準を線形的に結合する。

$c = \{x_1, \dots, x_M\}$ がラベルあり事例を含まない一つのクラスタで、 M 個の事例から構成されているとすると、事例 x_i の Representative 基準は、以下のように定義できる。

$$Rep(x_i) = \sum_{j \in c} w_{ij} \quad (4.6)$$

基本的には、ある事例のクラスタ内類似度が大きければ大きいほど、その事例はより良い Representative な事例である。

事例 x_i の Informative 基準は、Tong ら [9] と同様に、以下のように定義できる。

$$Inf(x_i) = |g(x_i)| \quad (4.7)$$

但し、 $g(x_i)$ は、SVM 分類器の実数値出力である。

これらの二つの基準を統合して、ラベルなし事例 x_i のスコアは、次式のように定義できる。

$$s(x_i) = \lambda Inf(x_i) + (1 - \lambda) Rep(x_i) \quad (4.8)$$

但し、 λ は、各要素の影響を調整するためのパラメータである。

$\lambda = 0$ の場合、提案手法は、絶対的に事例の Representative な性質に注目するが、 $\lambda = 1$ の場合には、 SVM_{Active} の Informative 基準の戦略に戻すことにはならない。 SVM_{Active} の Informative 基準は、すべてのラベルなし事例の中から最も Informative な事例を選択するが、提案手法の戦略は、個々のクラスタの中から一つの最も Informative な事例を選択するからである。

提案手法では、クラスタ c 内のすべての事例に対するスコアを式 (4.8) に基づいて計算し、最大のスコアを持つ事例を選択して、ユーザにラベリングを要求するために提示する。

4.3.4 ラベルなし事例の活用

ここでは，Co-Training の枠組みの下で，学習過程を強化し，検索性能を改善するために，ラベルなし事例の活用法を提案する．具体的には，SVM とは異なるもう一つの分類器を構成し，SVM の訓練事例を増やすために，そのもう一つの分類器のラベルなし事例に対する予測を使用する．

L と U をそれぞれラベルあり事例のデータセットとラベルなし事例のデータセットとすると，最初は， L はクエリであり， U はデータベース内のすべての事例である．適合フィードバックでは，ユーザが，いくつかの事例に対してクエリに適合 / 非適合のラベリングを実行する．この過程は，正例あるいは負例を追加することにより，ラベルあり事例のデータセットを増やすことに相当する．ラベルあり事例が増やされた L を用いて，もう一つの分類器が構成され， U 中の事例をそれらの予測に基づいてソートした後，この分類器によるいくつかの最も信頼できる負例を，SVM の訓練事例を増やすための負例として使用する．

提案手法では，CBIR の実時間処理の要求を充たすために，極めて単純な重み付きユークリッド距離モデルをもう一つの分類器として使用する．

$$f(x) = d(x, q) = \left(\sum_{i=1}^d (x^i - q^i)^2 * w^i \right)^{1/2} \quad (4.9)$$

但し， x と q は，それぞれ，分類されるべき事例，新しいクエリポイントであり， d 次元の特徴ベクトルである． w^i は各特徴ベクトルの要素に関連付けられた重みである．

P と N は，それぞれ，ラベルありの正例，負例で， $P \cap N = \emptyset$ ， $L = P \cup N$ とすると，最適なクエリポイントは次式で与えられる．

$$q = \frac{1}{|P|} \sum_{x_k \in P} x_k - \frac{1}{|N|} \sum_{x_k \in N} x_k \quad (4.10)$$

特徴ベクトルの要素の重み w^i は，各特徴ベクトルの要素の分布を反映するものである．

本手法でも Rui ら [48] と同様に，標準偏差に基づく重み計算のアプローチを採る．直観的には，すべての正例がある一つの特徴ベクトルの要素に対して類似の値をとるならば，この要素は，ユーザの検索意図を反映した良い指標となり，さまざまな値をとるならば，良い指標ではない．したがって，正例の一つの特徴ベクトルの要素のすべての値が与

えられると、それらの標準偏差の逆数がその重みの良い推定値となる。分散が小さいほど、その重みは大きくなり、逆もまた同様である。

$\{x_k, k = 1, \dots, |P|\}$ を正例の特徴ベクトル, $x_k^i, i = 1, \dots, d$ をその要素, σ^i を正例の特徴ベクトル集合 $\{x_k^i, k = 1, \dots, |P|\}$ の標準偏差とすると、次式を得る。

$$w^i = \frac{1}{\sigma^i} \quad (4.11)$$

正規化すると、次式となる。

$$w^i = \frac{1}{\sum_{i=1}^d w^i} \quad (4.12)$$

一つの正例だけしかない場合は、すべての特徴ベクトルの要素に対して同一の重みを採用する。

この重み付きユークリッド距離モデルに基づく分類器は強力ではなく、この分類器がラベルなし事例に割り当てるラベルは不正確である。この問題を軽減するために、Zhou[46]と類似した方法で、この分類器から得られる少数の最も確信度の高い負例のみを SVM の訓練事例を増やすために使用する。一般に、CBIR においては、あるクエリに対して適合するのは、データベース内のほんの少数の画像であり、大多数の画像は不適合となる。したがって、最も確信度の高い負例を訓練事例とするのは、より信頼できる方法である。また、この分類器によりラベリングされる負例は、SVM によって一時的に使用されるだけであり、フィードバックの各回では、増えたラベルあり事例に基づいて、新しい分類器が構成され、異なった負例が選択される。

4.4 提案手法の枠組み

図 4.2 に、提案する SVM の能動学習に基づく適合フィードバック手法の枠組みを示す。提案手法は、SVM 分類器、能動的事例選択器、ラベルなし事例生成器の三つの主要機能から成る。

フィードバックの各回において、まず、ユーザが能動的事例選択器によって提示される事例に対して、適合（正例）／不適合（負例）のラベリングを行うと、ラベリングされた事例が一時的に保存される。次に、一時的に保存されたラベルあり事例に基づいて、ラベルなし事例生成器がいくつかの負例を出力する。一時的に保存された負例とすべてのラベルあり事例とを用いて、訓練された SVM 分類器は、データベース内の事例を再評価し、検索結果を更新する。現在の推定に基づいて、能動的事例選択器は、ユーザに適合／非適

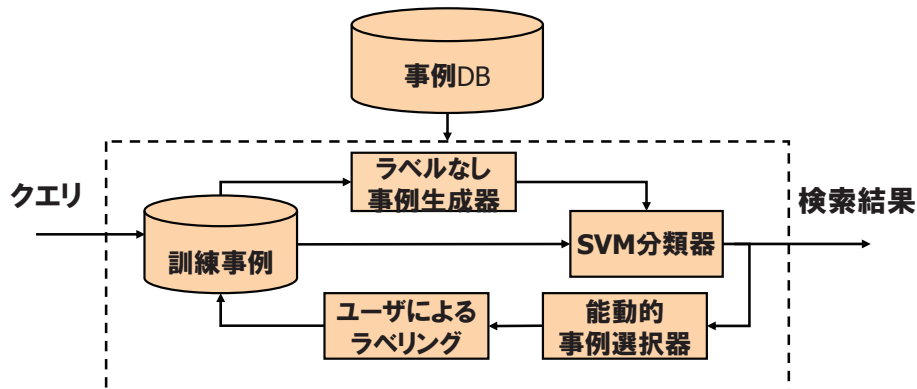


図 4.2 提案手法の枠組み

合のラベリングのために提示するいくつかの画像を選択する。ユーザが現在の検索結果に満足すれば、フィードバックは終了し、そうでなければ、ユーザは提示された事例に適合／非適合の評価のフィードバックを与え、次のフィードバックに入る。検索開始時点では、ユーザからのフィードバック情報は利用できず、クエリだけが正例として一時的に保存される。

要約すると、本論文で提案する SVM の能動学習に基づく適合フィードバック手法は、以下の通りである。

L と U を、それぞれ、ラベルあり事例のデータセット、ラベルなし事例のデータセットとする。

入力:

$query$: ユーザからのクエリ

DB : 事例データベース

初期化:

$L \leftarrow query$

$U \leftarrow DB$

フィードバック:

Step 1. L に基づいて、重み付きユークリッド距離分類器 $f(x)$ を構成する。

Step 2. $f(x)$ で U のすべての事例を予測し、 N^* で表す最も信頼できる負例を選択する。

Step 3. $L \cup N^*$ で、SVM を再訓練する。

Step 4. 現在の SVM で、 U のすべての事例を予測し、その予測値でソートすることにより、検索結果を出力する。

- Step 5. 現在の SVM の識別境界近傍の I_N で示す N 個の事例を選択し, $L \cup I_N$ に対してクラスタリングを実行する .
- Step 6. ラベルあり事例を含まないクラスから, L^* で表す Representative な事例を選択し, ユーザへのクエリとして提示する .
- Step 7. ユーザが現在の検索結果に満足すれば, フィードバックを終了し, そうでなければ, L^* に関するフィードバックを与える .
- Step 8. $L \leftarrow L \cup L^*, U \leftarrow U - L^*$ とし, Step 1. に戻る .

上記の手順から分かるように, 提案手法は事前に特別な事例へのラベリングを必要とせず, 検索開始時点では, ユーザはただ単に唯一の正例であるクエリを与えるだけで, SVM 分類器の訓練事例が自動的に生成される .

4.5 実験

提案手法の有効性を確認するために, CBIR に適用し, 評価実験を行った .

4.5.1 データセット

評価用データセットは, 以下の通りである .

組立図面の機械部品画像データセット

機械部品画像データセットは, 設計・製造分野において一般的に使用されている実際の組立図面から自動的に抽出した 14,770 枚の機械部品画像から成る . 組立図面は, 複数の部品とそれらの間の関係を記述しており, 例えば, ある製品やそのコンポーネントがどのような部品から構成され, どのように組み立てられているかを表現した図面である . また, このデータセットから, クエリ画像として使用するために, 組立図面で共通的に使用される 1,939 枚の機械部品画像を手動で選択し, 8 種類のカテゴリに分類した . カテゴリのラベルは, Ground Truth として, また, クエリ画像と同一のカテゴリに属する画像は, 正例の画像として使用する .

図 4.3 に, 典型的な機械部品画像を示す .

機械部品画像からは, 23 次元の Zernike モーメント特徴 [32], 128 次元の Grid Fourier 特徴 [29], [33], 108 次元のエッジ方向ヒストグラム特徴 [34] から構成される 259 次元の特徴ベクトルを抽出した .



図 4.3 機械部品画像の例

Corel 画像データセット

Corel 画像データセットは，鳥，海岸，風船，夕日などの 50 種類の概念カテゴリに分類された 5,000 枚の画像から成る．各概念カテゴリには 100 枚の画像が含まれており，各概念カテゴリは Ground Truth として使用する．

Corel 画像からは，64 次元の HSI 色ヒストグラム特徴 [19]，27 次元の Wavelet 変換係数特徴 [20]，[21] から構成される 91 次元の特徴ベクトルを抽出した．

4.5.2 実験内容

評価実験では，

- ラベルなし画像の活用
- 式 (4.8) のパラメータ λ
- ユーザにラベリングのために提示するフィードバック画像の数 L^*

が検索性能に及ぼす影響を調べるために，ラベルなしデータを活用する場合としない場合， λ および L^* をさまざまな値に設定した場合について，検索性能を評価した．

また，Brinker の Angle-Diversity 手法 [11] との検索性能の比較実験も行った．

検索性能の評価尺度

検索性能の評価尺度として，適合率 - 再現率曲線の BEP (Precision-Recall Break-Even Point) を使用した．BEP は，適合率 (Precision) と再現率 (Recall) が等しい点であり，クエリ画像に適合する評価用データセット中の画像の総数を T ，検索結果の上位 T 枚の画像群に含まれる適合画像数を R とすると，BEP の値は， R/T で与えられる．

実験方法

クエリ画像として、機械部品画像データセットについては 1,939 枚の画像を、Corel 画像データセットについては 5,000 枚の画像すべてを順番に与え、5 回の適合フィードバックを実行し、各回の BEP の値を計算した。

適合フィードバックの GUI

研究開発および評価実験を推進するために、適合フィードバックの GUI を開発した。

図 4.4～図 4.6 に、機械部品画像データセットに対する実験の画面例を示す。

各々の図において、左上部のウィンドウは、クエリ画像を、右上部のウィンドウは、検索結果を、右下部のウィンドウは、ユーザに適合 / 非適合の評価を要求するために選択された画像を提示するためのものである。右上部および右下部のウィンドウに提示された画像の下部には、“Yes (適合)”、“No (非適合)” のチェックボックスがあり、ユーザは、このボックスをチェックすることによって、適合 / 非適合の評価をシステムにフィードバックする。なお、ユーザは、右下部のウィンドウの画像については、必ず適合 / 非適合の評価を与えなければならないが、右上部の検索結果の画像については、必ずしも与える必要はない。

図 4.4 は、適合フィードバックの 0 回目、すなわち、ユーザがクエリ画像を与え、システムがそのクエリに対する検索結果とユーザに適合 / 非適合の評価を要求するために選択した画像を返した画面例である。これに対して、図 4.5 は、ユーザが適合 / 非適合の評価を与えた画面例であり、この評価がシステムにフィードバックされて、1 回の適合フィードバックを完了する。図 4.6 は、1 回の適合フィードバック後の画面例であり、検索性能が大きく改善されている様子が分かる。

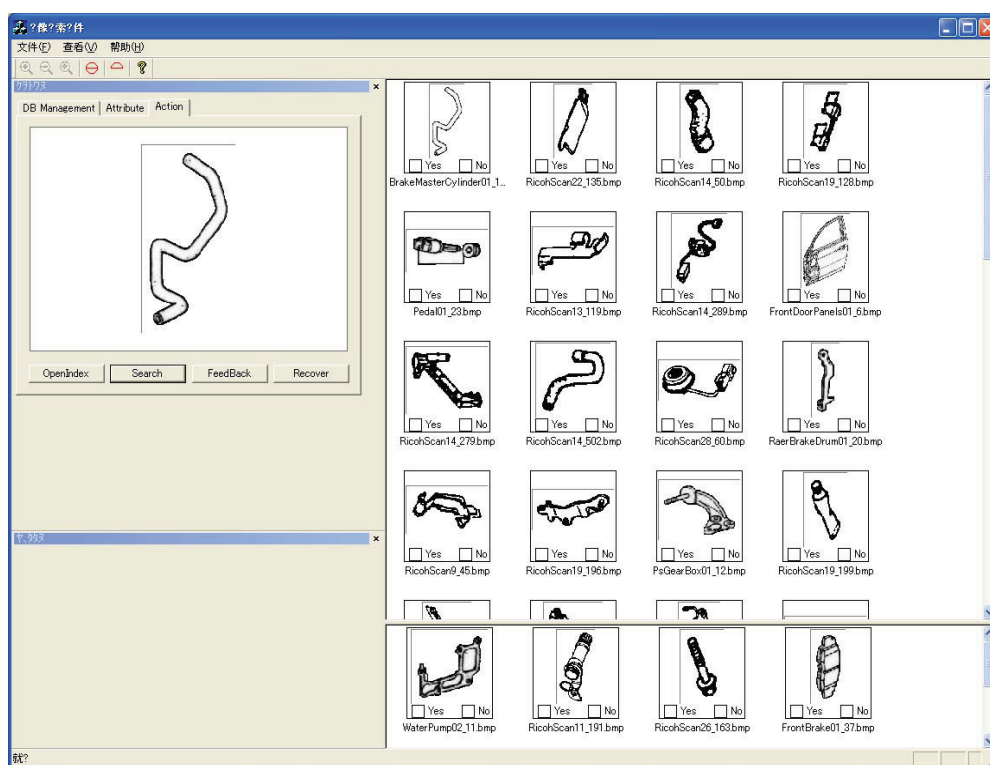


図 4.4 機械部品画像（反復回数=0）

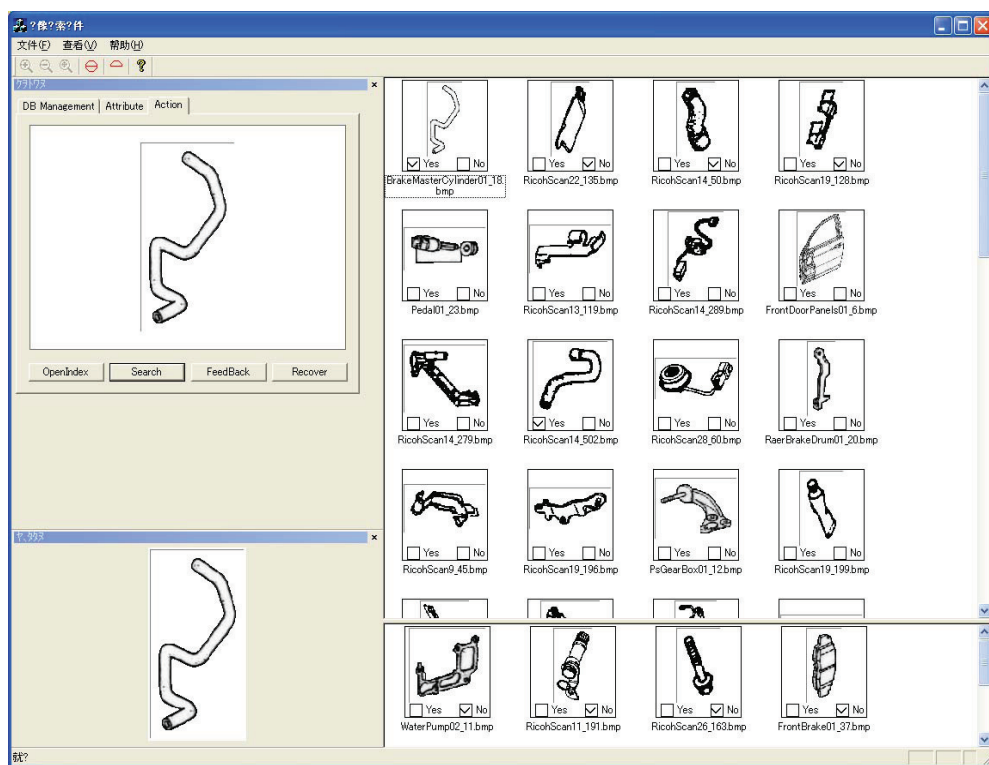


図 4.5 機械部品画像（反復回数=0，適合 / 非適合の評価）

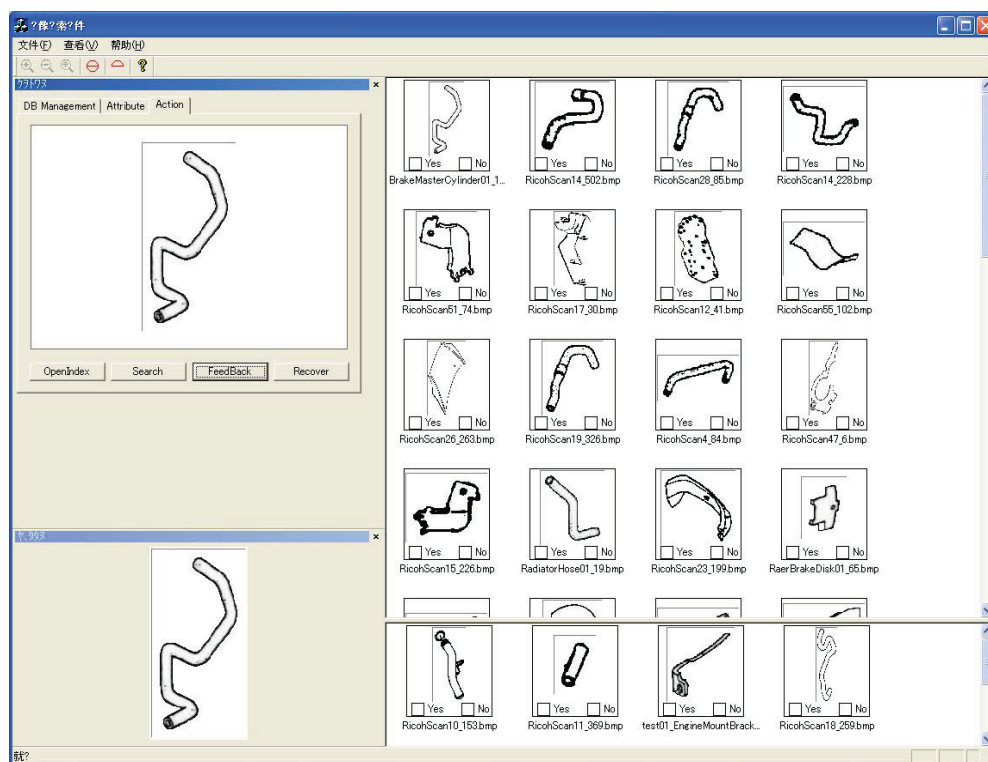


図 4.6 機械部品画像（反復回数=1）

4.5.3 結果と分析

ラベルなし画像の活用 / 非活用の影響

図 4.7 および図 4.8 に、10 枚のラベルなし画像を活用する場合とラベルなし画像を活用しない場合の BEP 曲線を示す。但し、ラベルなし画像を活用しない場合でも、SVM 分類器を正例、負例を用いて構成するために、0 回目のみ、ラベルなし画像を活用している。また、実験では、 $\lambda = 0.2, 0.5, 0.8$ 、 $L^* = 3, 6, 9$ とし、Normalized Cuts 法によるクラスタリングは、現在の SVM の識別境界近傍にある 400 枚の画像を対象として実行した。

図 4.7 および図 4.8 から、明らかに、ラベルなし画像の活用が検索性能の改善に有効であることが分かる。特に、 λ の値が大きく、フィードバック画像の数 L^* の値が小さい場合に、1 回目のフィードバックで大きな改善が見られている。

適合フィードバックでラベリングされる画像が少ない場合、すなわち、 L^* の値が小さい場合、通常、それらの画像が支持する特徴空間は、小さく極めて局所的な空間であり、大きいバイアスのかかった SVM 分類器が構成される。

同様に、 λ の値が大きい場合も、Informative 基準が Representative 基準よりも支配的になるため、画像間の冗長性が増し、それらが支持する特徴空間は小さくなる。

この場合、現在の SVM の識別境界から最も遠いという異なる観点から、最も信頼できる負例として得られるラベルなし画像は、SVM の識別超平面を大局的に調整するのを助け、汎化誤差を早く減少させることができる。

ラベルなし画像の活用の影響は、フィードバックの反復回数が進むにつれて、急激に小さくなる。最後の段階では、より多くの訓練事例が蓄積され、SVM 分類器の性能は急速に改善される。このような状態で、さらに性能改善を図るには、識別超平面を局所的に調整することが必要であるが、もう一つの分類器から負例として得られるラベルなし画像は、識別超平面から遠く離れた所にあり、性能改善にはほとんど寄与しない。

パラメータ λ の影響

図 4.9 に、パラメータ λ が検索性能に及ぼす影響を示す。ここでは、 $L^* = 6$ および $\lambda = 0.2, 0.5, 0.8, 1.0$ と設定し、10 枚のラベルなし画像を活用した。

二種類の評価用データセットに対する $\lambda = 0.5, 0.8$ の検索性能は非常に類似しており、一貫して $\lambda = 1.0$ よりも優れているが、小さい λ の値での検索性能はかなり異なっている。機械部品画像データセットでは、 $\lambda = 0.2$ の検索性能は、他の λ の値よりも多少優つ

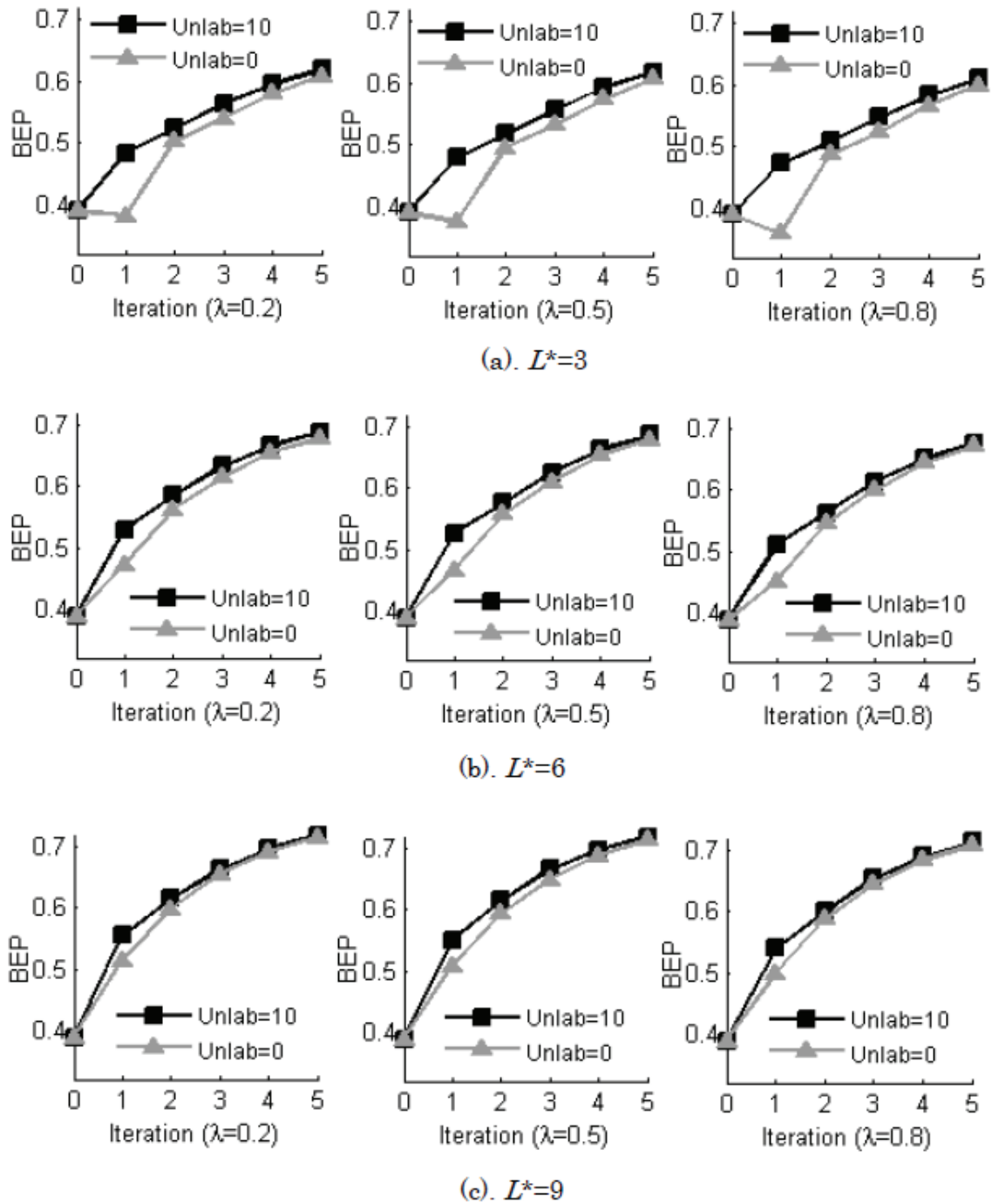


図 4.7 ラベルなし画像の活用 / 非活用 (機械部品画像)

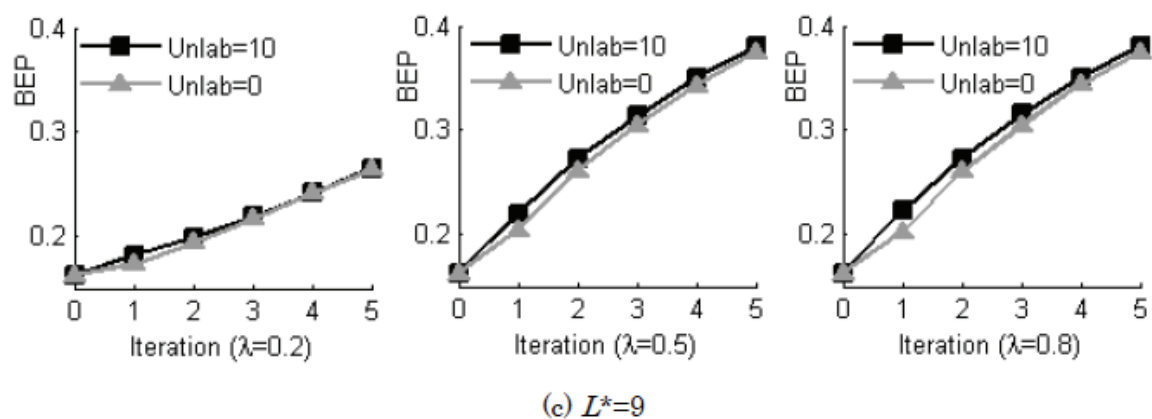
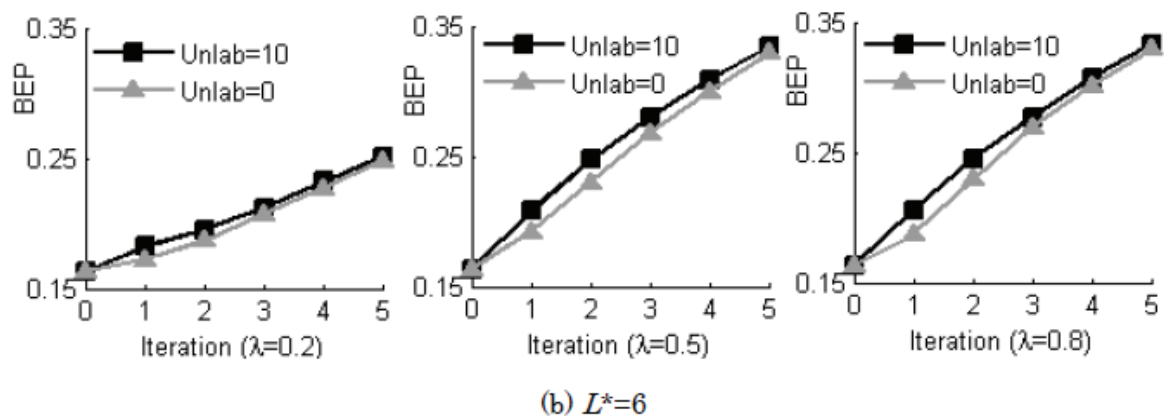
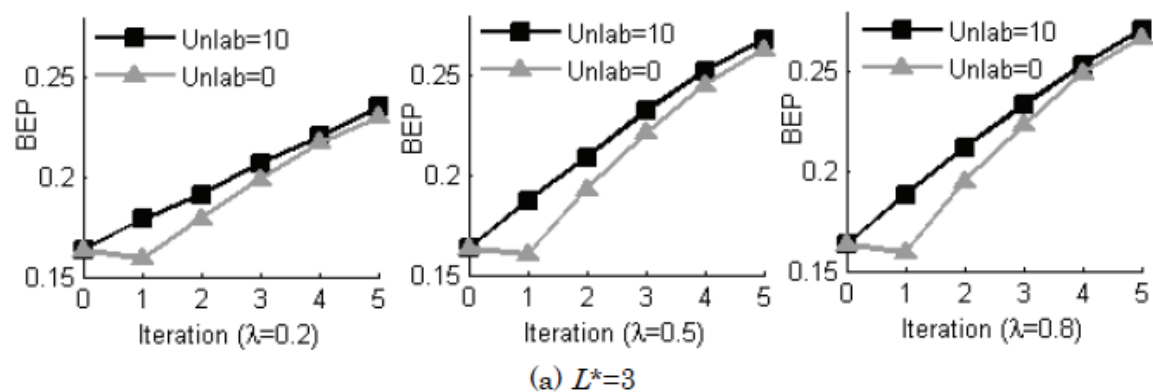


図 4.8 ラベルなし画像の活用 / 非活用 (Corel 画像)

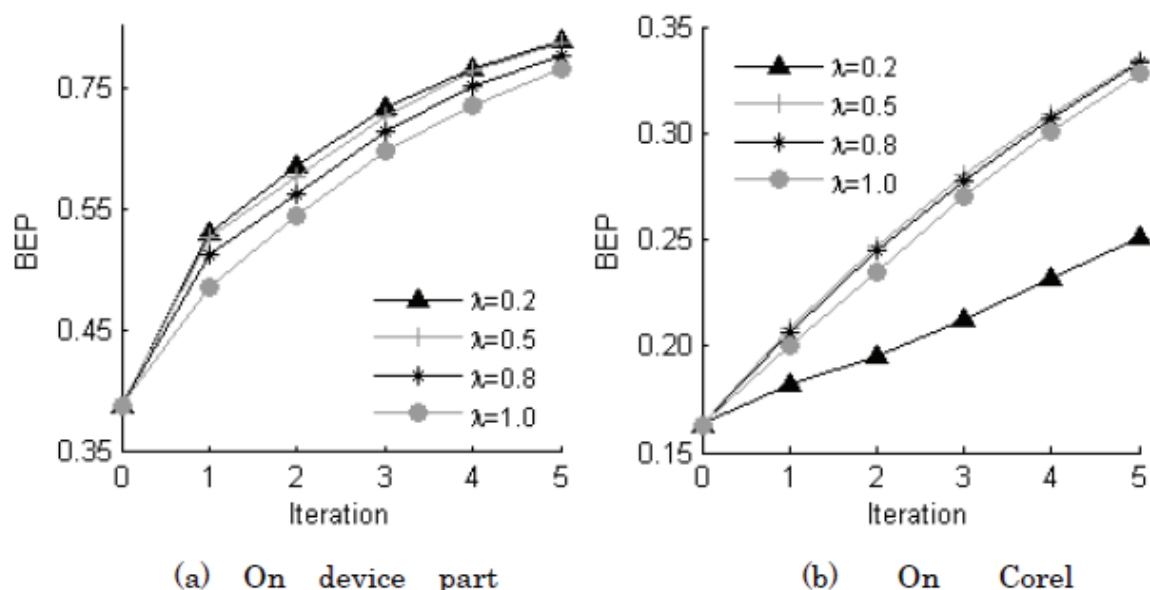
図 4.9 λ の影響 ($L^* = 6$)

表 4.1 Normalized Cuts 法による動的クラスタリングの実行時間

Times (s)	$L^* = 3$	$L^* = 6$	$L^* = 9$
Device part dataset	1.91	2.03	2.14
Corel dataset	1.47	1.50	1.61

ている．また，Representative 基準のみに基づく検索 ($\lambda = 0$) は，結合評価基準に基づく検索性能と匹敵しているが，Corel 画像データセットでは，小さい λ の値は，急激な性能低下を引き起こしている．

結論として， λ が検索性能に及ぼす影響は，評価用データセットにも多少依存するが， λ の値は 0.5～0.8 の範囲に設定するのが適切であると思われる．

Normalized Cuts 法による動的クラスタリングの実行時間は，グラフの構築に要する時間と Normalized Cuts 法の再帰的な適用に要する時間の和になる．グラフの構築には，重み行列 W の $N(N+1)/2$ の要素を評価する時間が必要であり，計算量のオーダーは $O(N^2)$ となる．なお， N はクラスタリングの対象とする画像数である．また，1回の Normalized Cuts 法の実行は，Lanczos アルゴリズムによって， $O(N^2)$ で解くことができるが，反復回数は L^* で決定される．実験では，現在の SVM 識別境界近傍にある 400 枚の画像が，Normalized Cuts 法によるクラスタリングの対象であり， $N = 400$ である．

表 4.1 に，2GB メモリ，3.06GHz の Pentium IV 搭載 PC での $L^* = 3, 6, 9$ での実行

時間を示す．

提案手法では，クラスタリングの実行時間に加えて，ラベルなし画像の生成，SVM の学習，予測のための実行時間が必要である．実験では，機械部品画像データセットに対しては数百ミリ秒，Corel 画像データセットに対しては数十ミリ秒程度の実行時間であった．したがって，1 回の適合フィードバックにかかる総実行時間は，機械部品画像データセットの場合が，約 2.5 秒，Corel 画像データセットが約 1.5 秒であり，実応用で十分に許容できるものであろう．

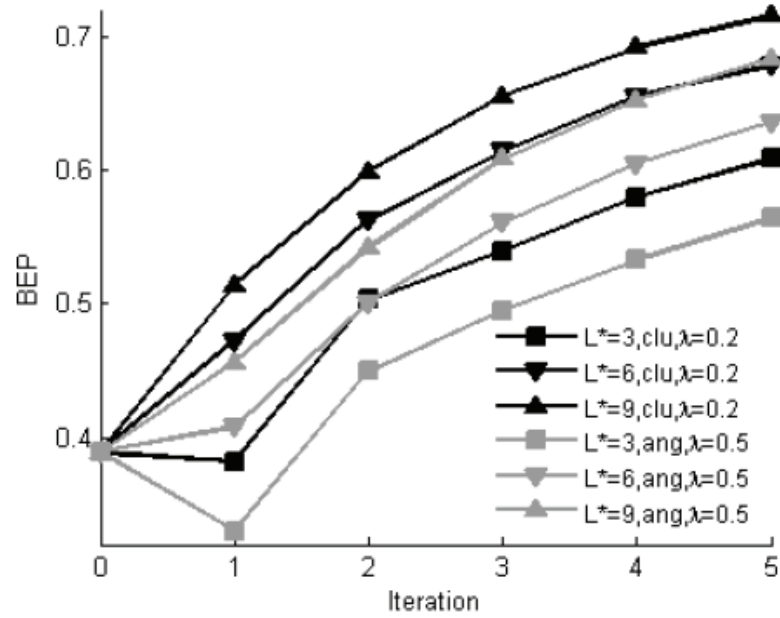
Angle-Diversity 手法との比較

本実験では， $L^* = 3, 6, 9$ の各々に対し，4.3.4 節で提案した活用法によって 10 枚のラベルなし画像を活用する場合と活用しない場合とで，提案手法と Angle-Diversity 手法 [11] との検索性能を比較した．また，各々の L^* に対して， $\lambda = 0.2, 0.5, 0.8$ と設定したときの検索性能を評価し， λ の最良の値を選択した．

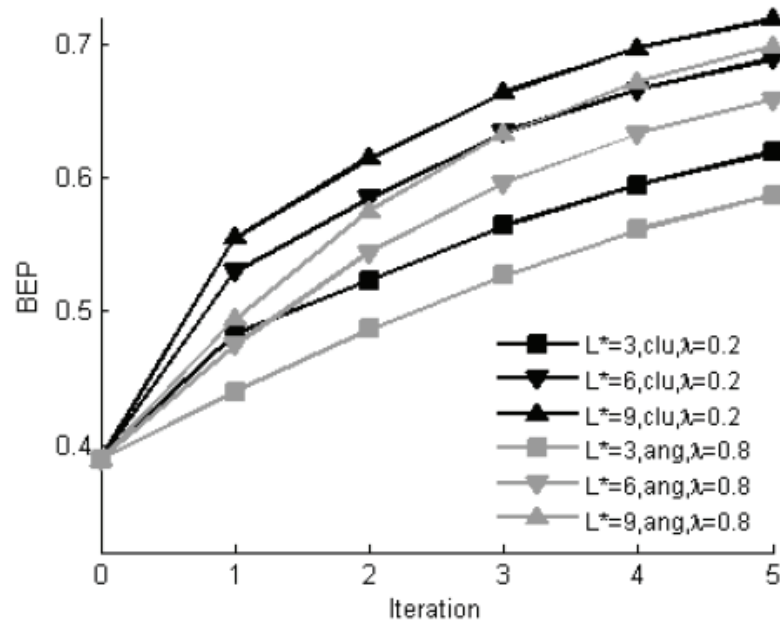
図 4.10 に，機械部品画像データセットを用いた実験結果を示す．提案手法では，ラベルなし画像を活用する場合もしない場合も， λ の最良の値は 0.2 であったが，Angle-Diversity 手法では，ラベルなし画像を活用する場合は 0.8，活用しない場合は 0.5 であった．

また，図 4.11 に，Corel 画像データセットを用いた実験結果を示す． λ の最良の値は，提案手法，Angle-Diversity 手法ともに，ラベルなし画像を活用するしないに拘らず，それぞれ 0.5，0.8 であった．

この実験結果から，明らかに，すべての回の適合フィードバックで，提案手法が SVM_{Active} Angle-Diversity 手法よりも優れており，特に，最初のフィードバックの段階で，より大きな性能改善を実現できることが分かる．

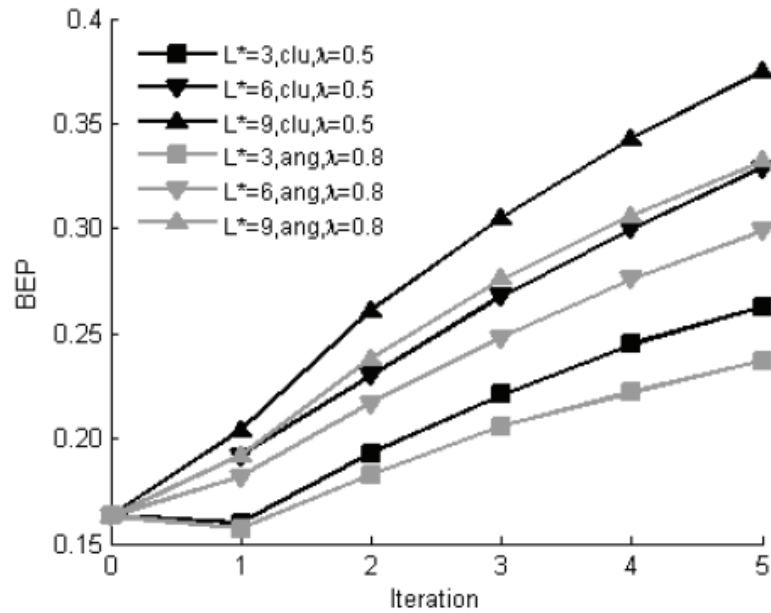


(a) Without unlabeled

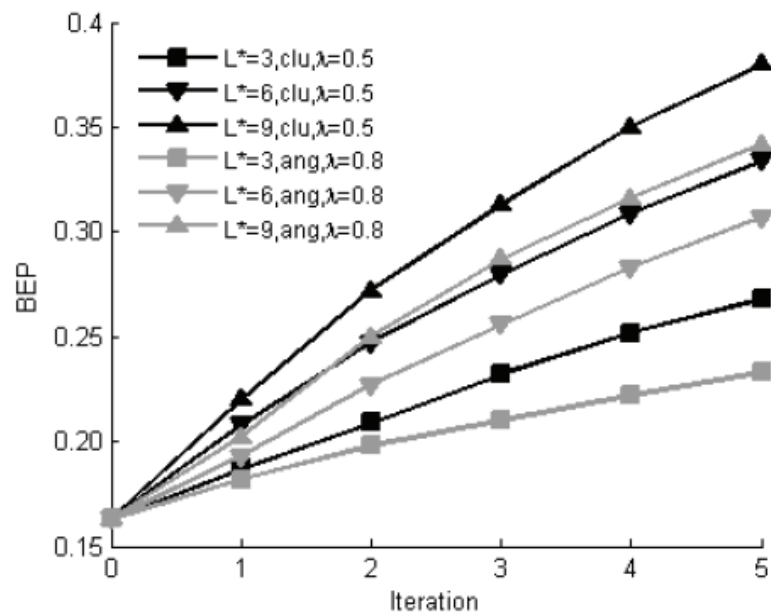


(b) With 10 unlabeled

図 4.10 Angle-Diversity 手法との比較 (機械部品画像)



(a) Without unlabeled



(b) With 10 unlabeled

図 4.11 Angle-Diversity 手法との比較 (Corel 画像)

4.6 まとめ

本章では，CBIR のための SVM の能動学習に基づく適合フィードバック手法を提案した．提案手法は，少ない訓練事例の問題を解決するために，以下の二つの戦略で，適合フィードバックの性能向上を実現した．

- 能動的事例選択の評価基準として，Representative 基準を設計し，Tong らの SVM_{Active} の Informative 基準と統合した新たな評価基準を提案した．この評価基準によると，選択される事例間の冗長性が減少し，より大きな情報ゲインが達成できる．
- Co-Training の枠組みの下で，ラベルなし事例の活用法を提案した．提案した活用法では，SVM の訓練事例を増やすために，SVM とは異なるもう一つの分類器を構成し，この分類器から得られる少数の最も信頼できる負例のみを訓練事例として活用する．これらのラベルなし事例は，一般に，SVM の識別超平面から遠くにあり，フィードバックの初期の段階で，大きな性能改善を実現できる．

これらの二つの戦略により，本提案手法は，SVM 分類器が生成する誤差を高速に減少し，ユーザの検索意図により良く適合することができる．

実際に，機械部品画像と Corel 画像を用いた CBIR に適用し，系統的な評価実験によって，本手法の有効性および優位性を実証した．

第 5 章

結論

5.1 本論文の主たる成果

本論文では、人間とシステムとの間に存在する Semantic Gap の解消を目指して、人間を積極的に検索および解析ループの中に取り込み、人間とシステムとのインタラクションを通して、互いの優れた能力を提供し、支援し合えるインタラクティブなマルチメディア・コンテンツの検索および解析手法を提案した。また、それらの手法を適用した実用システムを開発し、実応用面でも貢献した。

まず、2 章では、主に画像とテキストから構成されるマルチメディア・コンテンツを対象として、人間の視覚認識能力を最大限活用するためにシステムが何を支援すべきかを検討し、1) クロスメディア検索、2) 検索結果および検索過程の仮想三次元空間による視覚化、を特徴とするパノラミック検索モデルとそれに基づくパノラミック検索手法を提案した。本手法は、マルチメディア・コンテンツを構成する各メディアの特性およびメディア間の関連性を有効に活用し、コンテンツが画像とテキストで構成される場合には、画像による視覚的な検索とテキストによる意味的な検索を兼ね備えた情報検索手段を提供する。また、検索結果としてのコンテンツ集合やインタラクティブな検索過程を仮想三次元空間上で人間に分かりやすく視覚化することによって、人間の視覚認識能力を活用した効率的な検索が実現できる。本手法は、検索対象を“画像と関連テキスト”の対として表現する。検索対象とするマルチメディア・コンテンツがこの図式で表現できれば、本手法の枠組みがそのまま適用できる。その適用例として、Web 検索、映像検索、画像検索、デジタル文書検索、図面検索、3D-CAD モデル検索システムを開発し、実応用面でも貢献した。

3 章では、2 章で提案したパノラミック検索手法の枠組みを、大量のマルチメディア・コンテンツからの知識発見を支援する解析手法の枠組みへと拡張し、画像群からの知識発

見を支援するビジュアル解析手法として、パノラミック解析手法を提案した。画像群からの知識発見では、無数に定義できる視覚的特徴がある中で、有用な知識に関与する視覚的特徴を発見することが本質的な課題である。本手法は、人間を解析ループの中に取り込み、人間の高度な視覚認識能力を有効に活用し、人間による知識発見を支援する立場を採ることにより、インタラクティブに有用な知識に関与する視覚的特徴さらには有用な知識を発見できる。本手法を実装したパノラミック解析システムを開発し、発見学習などの教育分野や設計・製造分野での評価実験により、その分野の専門家でなくとも有用な知識の発見が可能であることを検証し、本手法の有効性を確認した。

4章では、人間とシステムとのインタラクションの一つとして、適合フィードバックの導入を検討し、1) 人間に適合 / 非適合の評価を求める事例を決定するための新しい Representative 基準、2) Co-Training の枠組みの下でのラベルなし事例の活用法、を提案した。また、それらを実装した SVM の能動学習に基づく適合フィードバック手法を開発した。本手法は、適合フィードバックの反復が少ない段階においても、人間の検索意図を反映した検索を効率的に実現できる。また、本手法を CBIR に適用した評価実験により、その有効性および従来手法に対する優位性を実証した。さらに、本手法を2章で提案したパノラミック検索手法に導入し、図面検索システムに適用することによって、実応用面でも貢献した。

5.2 今後の課題と展望

今後の課題としては、パノラミック検索および解析手法の機能強化、特に、システム側からの人間に対する支援機能の強化が挙げられる。人間を積極的に検索および解析ループの中に取り込むことで、Semantic Gap の解消を目指した本手法の枠組みの下でも、Semantic Gap の解消に向けた直接的なアプローチである視覚的特徴抽出機能の高度化は、取り組むべき重要な課題の一つである。この課題への取り組みの一つとして、画像中に存在する物体をそれが属する既知の概念カテゴリに分類する一般物体認識 (Generic Object Recognition) のアプローチがある。本手法にも、この一般物体認識のアプローチを導入し、より高次の意味概念レベルに近い視覚的特徴を抽出し、活用することができれば、Semantic Gap の解消に繋がり、より人間の検索意図を反映した検索、より高次の視覚的特徴が関与する有用な知識の発見が期待できる。この視覚的特徴抽出機能の高度化をはじめ、システム側からの人間に対する支援機能の強化という観点から、パノラミック検索および解析手法の機能強化、拡張を図っていく必要がある。手法の機能強化、拡張と並行して、さまざまな分野の実応用に展開していく予定である。

謝辞

本論文をまとめるにあたり，懇切なご指導と格別なご配慮をいただきました，東京大学大学院情報学環 池内克史教授に心より感謝申し上げます．また，東京大学生産技術研究所 上條俊介准教授，東京大学大学院情報学環 佐藤洋一准教授にはアドバイザ教員として研究の方向性についてご議論をいただきました．東京大学生産技術研究所 小野晋太郎特任助教，川上玲特任助教には的確なご助言をいただきました．深く感謝いたします．

論文の審査にあたり，適切なご教示と有益なご助言をいただきました東京大学大学院情報理工学系研究科 相澤清晴教授，安達淳教授，石塚満教授，東京大学生産技術研究所 喜連川優教授，および，豊田正史准教授に謹んで感謝申し上げます．

本研究の機会を与えていただきました株式会社富士通研究所 代表取締役社長 村野和雄氏，元常務取締役 林弘氏（現，独立行政法人 情報通信研究機構 監事），フェロー 浅川和雄氏に感謝の意を表します．また，本研究の初期段階で直接ご教示いただきました中京大学情報理工学部 棚橋純一教授（株式会社富士通研究所 元常務取締役）に心より感謝申し上げます．

共同研究を通して数多くの貴重なご討論と多大なご協力をいただいた株式会社富士通研究所 増本大器部長，上原祐介主任研究員，椎谷秀一研究員，遠藤進研究員，馬場孝之研究員，石原正樹研究員，指田直毅研究員，富士通研究開発中心有限公司 劉汝傑主任研究員，王月紅研究員には深く感謝いたします．また，実験を行うにあたり協力していただいた早稲田大学大学院 国際情報通信研究科 Chris Chu 氏にも感謝いたします．

九州大学大学院システム情報科学院 倉爪亮教授には本論文をまとめる段階で有益なご助言をいただきました．深く感謝いたします．

本論文をまとめるきっかけを与えてくださった東京工業大学名誉教授・日本ロボット学会名誉会長 梅谷陽二先生は，東京工業大学工学部機械物理工学科在籍時からの恩師であり，以来公私にわたりさまざまなご指導ご鞭撻をいただきました．感謝の念に堪えません．

最後に、これまで私をあたたく応援してくれた両親、私を明るく励まし続けてくれた妻、息子、娘に心から感謝します。

2009 年 6 月

発表文献

学術論文

1. 長田茂美, 増本大器, 上原祐介, 椎谷秀一, 遠藤進, 馬場孝之, 石原正樹, 池内克史, “パノラミック検索モデルに基づくインタラクティブなマルチメディア検索 / マイニングシステム “MIRACLES” ,” 電子情報通信学会論文誌 (D), Vol.J92-D, No.11, pp.1987-1999, 2009.
2. Rujie Liu, Yuehong Wang, Takayuki Baba, Daiki Masumoto, and Shigemi Nagata, “SVM-based active feedback in image retrieval using clustering and unlabeled data,” *Pattern Recognition*, Vol.41, No.8, pp.2645-2655, 2008.
3. 倉爪亮, 広瀬茂男, 岩崎倫三, 長田茂美, “協調ポジショニングシステムの研究 - CPS アクティブタッチ融合型地図生成法 - ,” 日本ロボット学会誌, Vol.17, No.1, pp.84-90, 1999.
4. 倉爪亮, 広瀬茂男, 岩崎倫三, 長田茂美, 指田直毅, “協調ポジショニングシステムの研究 - CPS-III を用いた自動清掃ロボットシステムの開発 - ,” 日本ロボット学会誌, Vol.16, No.7, pp.934-941, 1998.
5. 倉爪亮, 広瀬茂男, 岩崎倫三, 長田茂美, 指田直毅, “協調ポジショニングシステムの研究 - CPS-II 型システムの最適化移動形態 - ,” 日本ロボット学会誌, Vol.15, No.5, pp.773-780, 1997.
6. 倉爪亮, 広瀬茂男, 岩崎倫三, 長田茂美, 指田直毅, “協調ポジショニングシステムの研究 - 冗長位置情報の融合と機械モデル実験 - ,” 日本ロボット学会誌, Vol.14, No.8, pp.1229-1236, 1996.
7. 広瀬茂男, 倉爪亮, 長田茂美, “群ロボットによる協調ポジショニング法,” 日本ロボット学会誌, Vol.13, No.6, pp.838-845, 1995.
8. 長田茂美, 増本大器, 山川宏, 木本隆, “ニューラルネットによる感覚運動融合の階層型モデル,” 日本ロボット学会誌, Vol.12, No.5, pp.685-694, 1994.
9. Kazushige Saga, Tamami Sugasaka, Minoru Sekiguchi, Shigemi Nagata, and Kazuo Asakawa, “Mobile Robot Control by Neural Networks using Self-

- Supervised Learning," *IEEE Transactions on Industrial Electronics*, Vol.39, No.6, pp.537-542, 1992.
10. Shigemi Nagata, Minoru Sekiguchi, and Kazuo Asakawa, "Mobile Robot Control by a Structured Hierarchical Neural Network," *IEEE Control Systems Magazines*, Vol.10, No.3, pp.69-76, 1990.
 11. 関口実, 長田茂美, 浅川和雄, "構造化階層ニューラルネットワークによる移動ロボットの行動制御," *日本ロボット学会誌*, Vol.8, No.3, pp.66-74, 1990.
 12. 長田茂美, 渡部信雄, 浅川和雄, "ニューラルネットワークによる移動ロボットの制御," *電子情報通信学会論文誌 (D-II)*, Vol.J72-D-II, No.12, pp.2111-2120, 1989.
 13. Katsushi Ikeuchi, H. Keith Nishihara, Berthold K. P. Horn, Patrick Sobalvarro, and Shigemi Nagata, "Determining Grasp Configurations Using Photometric Stereo and The PRISM Binocular Stereo System," *International Journal of Robotics Research*, Vol.5, No.1, pp.46-65, 1986.
 14. 池内克史, 長田茂美, Berthold K. P. Horn, H. Keith Nishihara, "照度差ステレオとプリズムステレオによるピンピッキングタスク中のグリッパ姿勢決定," *電子通信学会論文誌 (D)*, Vol.J68-D, No.4, pp.546-553, 1985.
 15. 長田茂美, 井上彰, 吉田真澄, "論理回路図の自動入力処理," *電子通信学会論文誌 (D)*, Vol.J68-D, No.4, pp.837-844, 1985.

国際会議

1. Kenro Aihara, Taizo Yamada, Noriko Kando, Satoko Fujisawa, Yusuke Uehara, Takayuki Baba, Shigemi Nagata, Takashi Tojo, and Jun Adachi, "Supporting Creation and Sharing of Contents of Cultural Heritage Objects for Educational Purpose," *Proceedings of Workshop on Personalized Access to Cultural Heritage (PATCH 2008) in conjunction with 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2008)*, pp.2645-2655, 2008.
2. Yuehong Wang, Rujie Liu, Takayuki Baba, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "An Images Based 3D Model Retrieval Approach," *Proceedings of 14th International MultiMedia Modeling (MMM 2008)*, 2008.
3. Takayuki Baba, Susumu Endo, Shuichi Shiitani, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "An Interactive Image Mining System for Engineering Design and Manufacture," *Proceedings of Workshop on Knowledge Management and Semantic Web for Engineering Design (KW4ED) in conjunction with 4th International Conference on Knowledge Capture (ACM K-CAP 2007)*, 2007.

4. Noriko Kando, Kenro Aihara, Taizo Yamada, Satoko Fujisawa, Yusuke Uehara, Takayuki Baba, Shigemi Nagata, Takashi Tojo, Tetsuhiko Awaji, and Jun Adachi, "CEAX: A Flexible Content Management System for "Growing Metadata" of Cultural Heritage Objects and An Exploratory Search-and-Learn Mechanism for Guided Discovery," *Proceedings of International Workshop on Digital Preservation of Heritage and Research Issues in Archiving and Retrieval (IWDPH)*, 2007.
5. Rujie Liu, Yuehong Wang, Takayuki Baba, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "SVM-based Active Feedback in Image Retrieval using Clustering and Unlabeled Data," *Proceedings of 12th International Conference on Computer Analysis of Images and Patterns (CAIP 2007)*, 2007.
6. Rujie Liu, Takayuki Baba, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "Device Parts Retrieval from Assembly Drawings with SVM based Active Relevance Feedback," *Proceedings of ACM International Conference on Image and Video Retrieval (CIVR 2007)*, pp.379-386, 2007.
7. Susumu Endo, Takayuki Baba, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "Shape based 3D model retrieval without query," *Proceedings of ACM International Conference on Image and Video Retrieval (CIVR 2007)*, pp.218-225, 2007.
8. Kenro Aihara, Taizo Yamada, Noriko Kando, Satoko Fujisawa, Yusuke Uehara, Takayuki Baba, Shigemi Nagata, Takashi Tojo, and Jun Adachi, "Owlery: A Flexible Content Management System for "Growing Metadata" of Cultural Heritage Objects and Its Educational Use in the CEAX Project," *Proceedings of 9th International Conference on Asian Digital Libraries (ICADL 2006)*, pp.22-31, 2006.
9. Taizo Yamada, Kenro Aihara, Noriko Kando, Satoko Fujisawa, Yusuke Uehara, Takayuki Baba, Shigemi Nagata, Takashi Tojo, Yuko Hiroshima, and Jun Adachi, "CEAX's Learning Support System to Explore Cultural Heritage Objects without Keyword Search," *Proceedings of 9th International Conference on Multimedia and ICTs in Education (m-ICTE 2006)*, 2006.
10. Rujie Liu, Yu Hao, Takayuki Baba, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "Shape Similarity based on Contour Decomposition and Correspondence," *Proceedings of 8th International Conference on Signal Processing (ICSP 2006)*, pp.1015-1018, 2006.
11. Takayuki Baba, Rujie Liu, Susumu Endo, Shuichi Shiitani, Yusuke Uehara,

- Daiki Masumoto, and Shigemi Nagata, "Similarity-based Partial Image Retrieval System for Engineering Drawings," *Proceedings of 7th IEEE International Symposium on Multimedia (ISM 2005)*, pp.303-309, 2005.
12. Shuichi Shiitani, Takayuki Baba, Susumu Endo, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "Efficient Video Retrieval System Using Virtual 3D Space," *Proceedings of 6th IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.206-210, 2004.
 13. Susumu Endo, Shuichi Shiitani, Yusuke Uehara, Daiki Masumoto, and Shigemi Nagata, "MIRACLES: Multimedia Information Retrieval, Classification, and Exploration System," *Proceedings of International Conference on Multimedia and Expo 2002 (ICME 2002)*, Vol.2, pp.349-352, 2002.
 14. Yusuke Uehara, Susumu Endo, Shuichi Shiitani, Daiki Masumoto, and Shigemi Nagata, "A Computer-aided Visual Exploration System for Knowledge Discovery from Images," *Proceedings of 2nd International Workshop on Multimedia Data Mining (MDM/KDD 2001) in conjunction with 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.102-109, 2001.
 15. Naoshi Matsuo and Shigemi Nagata, "Study on Directional Microphone Technology using Estimated Signal," *Proceedings of ITC-CSCC '97*, pp.425-428, 1997.
 16. Naoki Sashida, Shigemi Nagata, and Ryo Kurazume, "Hierarchical Information Integration System using Bayesian Estimation," *Proceedings of 2nd Asian Conference on Computer Vision (ACCV '95)*, Vol.3, pp.77-81, 1995.
 17. Ryo Kurazume, Shigemi Nagata, and Shigeo Hirose, "Cooperative Positioning with Multiple Robots," *Proceedings of 1994 IEEE International Conference on Robotics and Automation*, Vol.2, pp.1250-1257, 1994.
 18. Daiki Masumoto, Hiroshi Yamakawa, Takashi Kimoto, and Shigemi Nagata, "Hierarchical Sensory-Motor Fusion Model with Neural Networks," *Proceedings of 1994 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI '94)*, pp.630-638, 1994.
 19. Hiroshi Yamakawa, Daiki Masumoto, Takashi Kimoto, and Shigemi Nagata, "Active Data Selection and Subsequent Revision for Sequential Learning with Neural Networks," *Proceeding of World Congress on Neural Networks (WCNN '94)*, Vol.3, pp.III/661-666, 1994.
 20. Takashi Kimoto, Daiki Masumoto, Hiroshi Yamakawa, and Shigemi Nagata, "Hierarchical Sensory Information Processing Model with Neural Networks," *Proceedings of 1993 IEEE International Conference on Robotics and Automation*,

- Vol.1, pp.929-934, 1993.
21. Daiki Masumoto, Takashi Kimoto, and Shigemi Nagata, "A Sensory Information Processing System using Neural Networks," *Proceedings of 1993 IEEE International Conference on Neural Networks (ICNN '93)*, Vol.2, pp.655-660, 1993.
 22. Minoru Sekiguchi, Tamami Sugasaka, and Shigemi Nagata, "Control of Multivariable System by a Neural Network," *Proceedings of 1991 IEEE International Conference on Robotics and Automation*, Vol.3, pp.2644-2649, 1991.
 23. Shigemi Nagata, Minoru Sekiguchi, Tamami Sugasaka, and Kazushige Saga, "Control of an Inverted Pendulum by a Neural Network with Self-Supervised Learning," *Proceedings of International Joint Conference on Neural Networks (IJCNN-91-Seattle)*, Vol.2, p.1003, 1991.
 24. Takashi Kimoto, Yoshinori Yaginuma, Shigemi Nagata, and Kazuo Asakawa, "Inverse Modeling of Dynamical System - Network Architecture with Identification Network and Adaptation Network -," *Proceedings of 1991 IEEE International Joint Conference on Neural Networks*, Vol.1, pp.566-571, 1991.
 25. Kazushige Saga, Tamami Sugasaka, and Shigemi Nagata, "Adaptive Behavior Control for a Robot by Neural Networks Using Self-supervised Learning," *Proceedings of 3rd International Workshop on Neural Networks and Their Applications (Neuro-Nimes '90)*, pp.433-441, 1990.
 26. Minoru Sekiguchi, Shigemi Nagata, and Kazuo Asakawa, "Behavior Control for a Mobile Robot by Multi-Hierarchical Neural Network," *Proceedings of 1989 IEEE International Conference on Robotics and Automation*, Vol.III, pp.1578-1583, 1989.
 27. Shigemi Nagata, Takashi Kimoto, and Kazuo Asakawa, "Control of Mobile Robots with Neural Networks," *Neural Networks*, Vol.1, No.1, SUPPL, p.349, 1988.
 28. Shinichi Shimizu, Shigemi Nagata, Akira Inoue, and Masumi Yoshida, "Logic Circuit Diagram Processing System," *Proceedings of 6th International Conference on Pattern Recognition*, Vol.2, pp.717-719, 1982.

研究会等

1. 石原正樹, 馬場孝之, 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, "異種メディア間の関連付けを利用したメディア横断検索システム," *映像情報メディア学会技術報告*, Vol.33, No.23, pp.73-76, 2009.

2. 椎谷秀一, 馬場孝之, 遠藤進, 上原祐介, 増本大器, 長田茂美, “マルチメディア情報検索システム「MIRACLES」の適用事例と今後の展望,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2007-34 / データ工学 DE2007-8, pp.43-48, 2007.
3. 馬場孝之, 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “画像群からの知識発見技術のものづくり分野への適用,” 情報処理学会研究報告, データベース・システム 2007-DBS-142 / 情報学基礎 2007-FI-87, pp.49-54, 2007.
4. 遠藤進, 馬場孝之, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “ブラウジングを用いた形状に基づく 3D モデル検索,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2007-17 / 医用画像 MI2007-17, pp.91-96, 2007.
5. 岩田則和, 長田茂美, 上原祐介, 馬場孝之, 津久間秀彦, 田中武志, 池内実, 石川澄, “時間軸を考慮した臨床検査データの空間配置の試行,” 医療情報学連合大会論文集, Vol.26, pp.546-549, 2006.
6. 山田太造, 相原健郎, 藤沢仁子, 神門典子, 上原祐介, 馬場孝之, 長田茂美, 安達淳, “学校教育における文化財コンテンツ利活用のための教育支援システム,” 日本教育工学研究会, 2006.
7. 馬場孝之, 劉汝傑, 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “機械系組み立て図面からの形状に基づく部品検索手法の提案,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2004-225 / 思考と言語 TL2004-57, pp.79-84, 2005.
8. 遠藤進, 馬場孝之, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “デジタル文書の共有・活用支援システム～マルチメディア検索システム MIRACLES による実現と評価～,” 情報処理学会研究報告, グループウェアとネットワークサービス 2004-GN-51, pp.91-96, 2004.
9. 上原祐介, 馬場孝之, 遠藤進, 椎谷秀一, 増本大器, 長田茂美, “仮想 3 次元空間での画像配置を利用したイメージマイニングの支援ツール,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2003-95, pp.1-6, 2003.
10. 椎谷秀一, 馬場孝之, 遠藤進, 上原祐介, 増本大器, 長田茂美, “画像情報とテキスト情報を統合的に利用したインタラクティブな映像検索システム,” 情報処理学会研究報告, データベース・システム 2002-DBS-127, pp.129-136, 2002.
11. 椎谷秀一, 遠藤進, 上原祐介, 増本大器, 長田茂美, “画像の色特徴を用いた商品検索サービス,” 情報処理学会研究報告, グループウェアとネットワークサービス 2002-GN-42, pp.37-42, 2002.
12. 長田茂美, 遠藤進, 椎谷秀一, 上原祐介, 増本大器, “マルチメディア情報検索システム “MIRACLES”,” 情報処理学会 人文科学とコンピュータシンポジウム 2001 論

文集, pp.267-274, 2001.

13. 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “テキストによる意味的な検索と画像による視覚的な検索を統合したマルチメディア検索システム“ MIRACLES ”,” *DBWeb 2001*, IPSJ Symposium Series Vol.2001, No.17, pp.249-256, 2001.
14. 上原祐介, 遠藤進, 椎谷秀一, 増本大器, 長田茂美, “仮想空間での情報構造表現に基づく画像群からの知識発見支援システム,” 人工知能学会研究会資料, 人工知能基礎論 / 知識ベースシステム SIG-FAI/KBS-J-40, pp.243-250, 2001.
15. 上原祐介, 遠藤進, 指田直毅, 増本大器, 長田茂美, 棚橋純一, “MIRACLES: マルチメディア情報のパノラミック検索システム - Web 検索への応用 -,” 電子情報通信学会技術研究報告, データ工学 DE2000-3, pp.17-24, 2000.
16. 遠藤進, 指田直毅, 増本大器, 長田茂美, 棚橋純一, “画像情報とテキスト情報を統合的に利用したインタラクティブな Web 検索システム,” 第 5 回知能情報メディアシンポジウム論文集, pp.163-170, 1999.
17. 北川博紀, 松尾直司, 原田裕明, 長田茂美, “カメラとマイクアレイの相互位置関係を利用した話者位置検出システム,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU98-79, pp.43-50, 1998.
18. 松尾直司, 長田茂美, “推定信号を用いたマイクロホンアレイの検討,” 電子情報通信学会技術研究報告, 応用音響 EA97-106, pp.25-31, 1998.
19. 指田直毅, 増本大器, 長田茂美, 池内克史, “Eigen-Window 法を用いた顔画像に基づく個人識別,” 第 15 回日本ロボット学会学術講演会予稿集, Vol.2, pp.361-362, 1997.
20. 長田茂美, 指田直毅, 倉爪亮, “ベイズ推定を用いた階層型情報統合システムの提案,” 電子情報通信学会 1995 年総合大会講演論文集, 情報・システム (2), p.339, 1995.
21. 山川宏, 増本大器, 木本隆, 長田茂美, “逐次学習における学習データの選択と追加的出力補正,” 電子情報通信学会技術研究報告, ニューロコンピューティング NC92-99, pp.33-40, 1993.
22. 関口実, 菅坂玉美, 長田茂美, “構造化ニューラルネットを用いた倒立振子の制御,” 日本機械学会ロボティクス・メカトロニクス講演会'92 論文集, Vol.A, pp.857-862, 1992.
23. 菅坂玉美, 関口実, 長田茂美, “ニューラルネットワークによる倒立振子の振り上げと倒立制御,” 第 17 回システムシンポジウム / 第 14 回知能システムシンポジウム / 第 1 回ニューラルネットワークシンポジウム合同シンポジウム論文集, pp.63-67, 1991.
24. 関口実, 菅坂玉美, 長田茂美, “ニューラルネットワークによる非線形多入出力系の制御,” 第 8 回日本ロボット学会学術講演会予稿集, pp.869-872, 1990.
25. 関口実, 長田茂美, “ニューラルネットワークによるロボットマニピュレータの逆

- ダイナミクス制御,” 第 2 回ロボティクス・自動化システムシンポジウム論文集, pp.185-189, 1990.
26. 長田茂美, 渡部信雄, 浅川和雄, “ニューラル・ネットによるロボット制御 -高速学習アルゴリズムと構造化ネットワーク-,” 電子情報通信学会技術研究報告, 集積回路 ICD88-126, pp.57-64, 1988.
27. 長田茂美, 関口実, 吉沢英樹, 渡部信雄, 木本隆, 浅川和雄, “ニューロコンピュータによるロボットの行動制御,” 情報処理学会研究報告, 情報学基礎 FI11-2, 1988.
28. 木本隆, 長田茂美, 関口実, 吉沢英樹, 渡部信雄, 浅川和雄, “ニューラル・ネットワークによる移動ロボットの制御,” 電子情報通信学会技術研究報告, 計算機アーキテクチャ ARC71-14, pp.101-108, 1988.
29. 木本隆, 長田茂美, 関口実, 吉沢英樹, 渡部信雄, 浅川和雄, “ニューラル・ネットワークの学習方式の検討,” 1988 年度人工知能学会全国大会論文集 3-7, pp.123-126, 1988.
30. 長田茂美, 関口実, 吉沢英樹, 渡部信雄, 木本隆, 浅川和雄, “ニューラル・ネットワークによる移動ロボットの制御,” 1988 年度人工知能学会全国大会論文集 3-8, pp.127-130, 1988.
31. 松浦俊夫, 長田茂美, 岩田清, 吉田真澄, “論理回路図の自動入力装置,” 情報処理学会研究報告, コンピュータビジョン CV31-1, 1984.
32. 長田茂美, 池内克史, “MIT におけるコンピュータビジョン研究の現状,” 情報処理学会研究報告, コンピュータビジョン CV28-6, 1984.
33. 池内克史, 長田茂美, ベソルド ホーン, キース ニシハラ, “ピンピッキングのための目,” 電子通信学会技術研究報告, パターン認識と学習 PRL83-53, pp.69-84, 1983.
34. 長田茂美, 井上彰, 吉田真澄, “線画像認識方式 - 文字と図形が混在する図面 -,” 情報処理学会研究報告, コンピュータビジョン CV17-3, 1982.
35. 清水慎一, 榊井猛, 長田茂美, 吉田真澄, 織田勇, “設計図面の自動入力 / 処理装置 - 格子軸上に手書きされた図形 -,” 電子通信学会技術研究報告, パターン認識と学習 PRL80-53, pp.97-104, 1980.
36. 長田茂美, 井上彰, 榊井猛, 松浦俊夫, 吉田真澄, “線図形の読み取り方式 - 格子軸上に手書きされた図形 -,” 電子通信学会技術研究報告, パターン認識と学習 PRL80-32, pp.39-46, 1980.
37. 長田茂美, 吉田真澄, 岩田清, 榊井猛, “線図形の記述方式,” 電子通信学会技術研究報告, パターン認識と学習 PRL79-50, pp.27-36, 1979.

著書・編著書（分担執筆）

1. 電子情報通信学会編, “デジタル信号処理ハンドブック,” オーム社, 1993.
2. 科学技術庁監修, 山崎弘郎, 石川正俊編著, “センサフュージョン -実世界の能動的理解と知的再構成-,” コロナ社, 1992.
3. 岡部洋一監修, “事例にみるニューラルネットワークス応用技術,” トリケップス社, 1992.
4. 日本機械学会編, “絵で学ぶバイオテクノロジー入門,” オーム社, 1991.
5. 北村新三, 片井修編, “画像と制御 -ニューロ・ファジィ・AI によるアプローチ-,” 朝倉書店, 1991.

総説・学術資料等

1. 増本大器, 馬場孝之, 遠藤進, 椎谷秀一, 上原祐介, 長田茂美, “情報を眺めて選ぶマルチメディア検索システム MIRACLES,” 情報の科学と技術, Vol.54, No.11, pp.582-588, 2004.
2. Naoshi Matsuo, Hiroki Kitagawa, and Shigemi Nagata, “Speaker Position Detection System Using Audio-visual Information,” *Fujitsu Scientific and Technical Journal*, Vol.35, No.2, pp.212-220, 1999.
3. 松尾直司, 北川博紀, 長田茂美, “マイクロホンアレイを用いた音声入力インタフェース,” *FUJITSU*, Vol.49, No.1, pp.80-84, 1998.
4. 長田茂美, “ニューラルネット: 技術と応用の動向,” 計量管理, Vol.39, No.1, pp.32-39, 1990.
5. 長田茂美, “ニューロコンピュータの原理と応用,” 1990 年度人工知能学会全国大会 (第4回) チュートリアル資料, 1990.
6. 長田茂美, “ニューラル・ネットワークによる移動ロボットの制御,” 計測自動制御学会「ロボットやビジョンの知能化に関する特別講演会 (第2回知識工学部会)」資料, 1989.
7. 長田茂美, 木本隆, 浅川和雄, “ニューロコンピュータの原理とロボット制御への応用,” *FUJITSU*, Vol.39, No.3, pp.175-184, 1988.
8. Shigemi Nagata, Toshio Matsuura, and Hideichi Endo, “Automatic Recognition System for Logic Circuit Diagrams,” *FUJITSU Scientific and Technical Journal*, Vol.21, No.4, pp.408-420, 1985.
9. 吉田真澄, 梶井猛, 長田茂美, 織田勇, “パターン認識技術の応用: 形状認識技術への応用 (3): 手書図面の自動入力 / 処理装置,” 情報処理, Vol.22, No.4, 1981.

参考文献

- [1] M. S. Kankanhalli and Y. Rui, "Application Potential of Multimedia Information Retrieval," *Proceedings of the IEEE*, Vol.96, No.4, pp.712-720, 2008.
- [2] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-Based Multimedia Information Retrieval: State of the Art and Challenges," *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol.2, No.1, pp.1-19, 2006.
- [3] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys*, Vol.40, No.2, article 5, pp.5-1-5-60, 2008.
- [4] 串間和彦, 赤間浩樹, 紺谷精一, 山室雅司, "色や形状等の表層的特徴量に基づく画像内容検索技術," *情報処理学会論文誌*, Vol.40, No.SIG3(TOD1), pp.171-184, 1999.
- [5] T. Kohonen, S. Kaski, K. Lagus, J. Salojärvi, J. Honkela, V. Paatero, and A. Saarela, "Self Organization of a Massive Document Collection," *IEEE Transactions on Neural Networks*, Vol.11, No.3, pp.574-585, 2000.
- [6] 武者義則, 広池敦, 杉本晃宏, "類似画像検索における特徴量空間の可視化インタフェース," *電子情報通信学会論文誌 (D-II)*, Vol.J82-D-II, No.10, pp.1626-1633, 1999.
- [7] 木本晴夫, "3次元探索空間を用いるインタラクティブな画像検索システムとその検索特性," *情報処理学会論文誌*, Vol.41, No.3, pp.638-657, 2000.
- [8] ImageCruiser Labs., ImageCruiser, <http://imagecruiser.jp/>.
- [9] S. Tong and E. Chang, "Support Vector Machine Active Learning for Image Retrieval," *Proceedings of 9th ACM International Conference on Multimedia*, pp.107-118, 2001.
- [10] J. He, M. Li, H. J. Zhang, H. Tong, and C. Zhang, "Mean Version Space: a New Active Learning Method for Content-Based Image Retrieval," *Proceedings of 6th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR)*, pp.15-22, 2004.
- [11] K. Brinker, "Incorporating Diversity in Active Learning with Support Vector

- Machines," *Proceedings of 20th International Conference on Machine Learning (ICML 2003)*, pp.59-66, 2003.
- [12] 月本洋, 柿元満, 森田千絵, 菊池吉晃, "脳機能画像のノンパラメトリック回帰分析," 電子情報通信学会論文誌 (D-II), Vol.J84-D-II, No.12, pp.2623-2633, 2001.
- [13] P. Stolorz, H. Nakamura, E. Mesrobian, R. R. Muntz, E. C. Shek, J. R. Santos, J. Yi, K. Ng, S. Y. Chien, C. R. Mechoso, and J. D. Farrara, "Fast Spatio-Temporal Data Mining of Large Geophysical Datasets," *Proceedings of 1st International Conference on Knowledge Discovery and Data Mining*, pp.300-305, 1995.
- [14] 片山幸司, 小西修, "知識発見支援のための気象画像データベースの構築," 情報処理学会論文誌, Vol.40, No.SIG5(TOD2), pp.69-78, 1999.
- [15] A. Kitamoto, "Data Mining for Typhoon Image Collection," *Proceedings of 2nd International Workshop on Multimedia Data Mining (MDM/KDD 2001) in conjunction with 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.68-77, 2001.
- [16] O. R. Zaiane, M. L. Antonie, and A. Coman, "Mammography Classification by an Association Rule-based Classifier," *Proceedings of 3rd International Workshop on Multimedia Data Mining (MDM/KDD 2002) in conjunction with 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.62-69, 2002.
- [17] J. Tesic, S. Newsam, and B. S. Manjunath, "Mining Image Datasets using Perceptual Association Rules," *Proceedings of SIAM 6th Workshop on Mining Scientific and Engineering Datasets in conjunction with 3rd SIAM International Conference (SDM)*, 2003.
- [18] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," *IEEE Computer*, Vol.28, No.9, pp.23-32, 1995.
- [19] 高木幹雄, 下田陽久監修, 新編 画像解析ハンドブック, 東京大学出版会, 2004.
- [20] W. Y. Ma and B. S. Manjunath, "A Comparison of Wavelet Transform Features for Texture Image Annotation," *Proceedings of International Conference on Image Processing*, pp.256-259, 1995.
- [21] 村尾晃平, 安藤淳禎, "画像をキーとする類似画像検索システム," 1998 年電子情報通信学会 情報・システムソサイエティ大会論文集, D-11-60, p.175, 1998.
- [22] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, Vol.41, No.6, pp.391-407, 1990.
- [23] 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, "テキストによる意味的な検索と

- 画像による視覚的な検索を統合したマルチメディア検索システム “MIRACLES”,
DBWeb 2001, IPSJ Symposium Series Vol.2001, No.17, pp.249-256, 2001.
- [24] 山根康男, 三末和男, “規則的球配置と方向に基づく近似を特徴とする多次元インデックス方式,” 情報処理学会研究報告, データベース・システム 2002-DBS-127, pp.65-72, 2002.
- [25] T. Kohonen (徳高平蔵他 訳), 自己組織化マップ, シュプリンガー・フェアラーク東京, 2005.
- [26] G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, NewYork, 1983.
- [27] 株式会社富士通九州システムズ, みよう絵, <http://jp.fujitsu.com/group/kyushu/services/imaging/miyoue/>.
- [28] 遠藤進, 馬場孝之, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “デジタル文書の共有・活用支援システム～マルチメディア検索システム MIRACLES による実現と評価～,” 情報処理学会研究報告, グループウェアとネットワークサービス 2004-GN-51, pp.91-96, 2004.
- [29] 馬場孝之, 劉汝傑, 遠藤進, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “機械系組み立て図面からの形状に基づく部品検索手法の提案,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2004-225 / 思考と言語 TL2004-57, pp.79-84, 2005.
- [30] 遠藤進, 馬場孝之, 椎谷秀一, 上原祐介, 増本大器, 長田茂美, “ブラウジングを用いた形状に基づく 3D モデル検索,” 電子情報通信学会技術研究報告, パターン認識・メディア理解 PRMU2007-17 / 医用画像 MI2007-17, pp.91-96, 2007.
- [31] Y. Wang, R. Liu, T. Baba, Y. Uehara, D. Masumoto, and S. Nagata, “An Images Based 3D Model Retrieval Approach,” *Proceedings of 14th International MultiMedia Modeling (MMM 2008)*, 2008.
- [32] W. Y. Kim and Y. S. Kim, “A region-based shape descriptor using Zernike moment,” *Signal Processing: Image Communication*, Vol.16, No.1-2, pp.95-102, 2000.
- [33] M. Safar, C. Shahabi, and X. Sun, “Image Retrieval by Shape: A Comparative Study,” *Proceedings of IEEE International Conference on Multimedia and Expo*, Vol.1, pp.141-144, 2000.
- [34] R. Liu, T. Baba, Y. Uehara, D. Masumoto, and S. Nagata, “Device Parts Retrieval from Assembly Drawings with SVM based Active Relevance Feedback,” *Proceedings of ACM International Conference on Image and Video Retrieval (CIVR 2007)*, pp.379-386, 2007.
- [35] J. W. H. Tangelder and R. C. Veltkamp, “A Survey of Content Based 3D Shape

- Retrieval Methods," *Proceedings of International Conference on Shape Modeling*, pp.145-156, 2004.
- [36] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape Distribution," *ACM Transactions on Graphics*, Vol.21, No.4, pp.807-832, 2002.
- [37] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. Kunii, "Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes," *Proceedings of SIGGRAPH 2001*, pp.203-212, 2001.
- [38] 桑原明裕, 小山聡, 角谷和俊, 田中克己, "マルチメディア・メタサーチのための質問変換と検索結果の統合," *DBSJ Letters*, Vol.2, No.1, pp.55-58, 2003.
- [39] Google Inc., Google Image Search, <http://images.google.co.jp/>.
- [40] アクセラテクノロジ株式会社, Accela BizSearch, <http://www.accelatech.com/products/BS/>.
- [41] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rules between Sets of Items in Large Databases," *Proceedings of ACM SIGMOD Conference on Management of Data*, pp.207-216, 1993.
- [42] P. Langley, "The Computer-Aided Discovery of Scientific Knowledge," *Proceedings of 1st International Conference on Discovery Science*, pp.25-39, 1998.
- [43] K. Aihara, T. Yamada, N. Kando, S. Fujisawa, Y. Uehara, T. Baba, S. Nagata, T. Tojo, and J. Adachi, "Supporting Creation and Sharing of Contents of Cultural Heritage Objects for Educational Purpose," *Proceedings of Workshop on Personalized Access to Cultural Heritage (PATCH 2008) in conjunction with 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2008)*, pp.2645-2655, 2008.
- [44] N. Kando, K. Aihara, T. Yamada, S. Fujisawa, Y. Uehara, T. Baba, S. Nagata, T. Tojo, T. Awaji, and J. Adachi, "CEAX: A Flexible Content Management System for "Growing Metadata" of Cultural Heritage Objects and An Exploratory Search-and-Learn Mechanism for Guided Discovery," *Proceedings of International Workshop on Digital Preservation of Heritage and Research Issues in Archiving and Retrieval (IWDPH)*, 2007.
- [45] K. Aihara, T. Yamada, N. Kando, S. Fujisawa, Y. Uehara, T. Baba, S. Nagata, T. Tojo, and J. Adachi, "Owlery: A Flexible Content Management System for "Growing Metadata" of Cultural Heritage Objects and Its Educational Use in the CEAX Project," *Proceedings of 9th International Conference on Asian Digital Libraries (ICADL 2006)*, pp.22-31, 2006.
- [46] Z. H. Zhou, K. J. Chen, and H. B. Dai, "Enhancing Relevance Feedback in Im-

- age Retrieval using Unlabeled Data," *ACM Transactions on Information Systems*, Vol.24, No.2, pp.219-244, 2006.
- [47] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp.888-905, 2000.
- [48] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.8, No.5, pp.644-655, 1998.

