特別研究報告書

# ユーザの画像間の違いの感じ方を利用する類似画像検索システム

指導教官　　池田 克夫 教授

京都大学工学部情報学科

上田 岳史

平成 12 年 2 月 15 日

# ユーザの画像間の違いの感じ方を利用する類似画像検索システム

上田 岳史

内容梗概

本研究では，検索キーとして画像を用いる類似画像検索の一手法である画像ブラウジングによる検索において，ユーザの入力からユーザの画像間の違いの感じ方を推定してそれを利用する画像検索システムを提案する。

近年の計算機の性能向上や記憶装置の発達などにより，画像データを大量に蓄積することが可能になった。またその用途も多様化しており，蓄積された画像のなかからユーザが希望するものを効率よく，かつ簡単に検索できる技術が望まれている。

その画像検索技術の一つとして，画像ブラウジングによる検索がある。画像ブラウジングによる検索とは，システムの数個から 10 数個の選択用画像の表示とユーザの選択の繰り返しによって候補を絞りこんでいく手法である。この手法には，ユーザが希望する画像についてはっきりとしたイメージを持っていなくても検索が可能で，また選択操作のみでよいことから，簡単に扱うことができるという利点がある。

従来の画像ブラウジングによる検索には，前もってデータベース画像をある基準に基づいてクラスタリングしユーザにクラスタの代表画像を選択させる手法と，ユーザの選択した画像とデータベース画像を直接比較する手法の二つがある。しかし，これらの手法にはシステムがどのように候補を絞りこむのかがユーザに分かりにくい，検索結果の画像が選択用画像を兼ねているためユーザが自分のイメージに合う画像を選択しづらいという二つの問題点がある。

そこで本研究では，以上の問題点を解決するための候補の絞りこみ手法とインタフェースを提案する。特徴量空間内にはユーザの希望する画像のクラスタと，ユーザがその画像とはある観点で異なると感じる画像のクラスタが存在し，ユーザはこれらのクラスタ間で違いを顕著に感じ，それらを結ぶ直線がユーザの画像間の違いの感じ方を反映するものと仮定する。この仮定に基づき，ユーザの画像間の違いの感じ方を反映する特徴量空間内の直線を，ユーザの希望する画像から近いことを表す○ ，遠いことを表す× という二種類の入力から推定

する。そして絞りこみ手法として，二つのクラスタを結ぶ直線上に境界点をとりその直線上に画像を射影しユーザが希望する画像とは異なると感じるクラスタ側の画像を検索範囲から取り除く方法をとる。ユーザの画像間の違いの感じ方を反映した直線上で候補を絞りこむことにより，ユーザに絞りこみの様子が分かりやすくなる。また，インタフェースとして，ユーザの入力を受け付ける選択用画像と結果表示用の候補画像を分離してユーザに提示するインタフェースを提案する。選択用画像に表示される画像として，推定した直線上に画像を射影したときに離れた位置に射影される画像を用いることにより，選択用画像に表示される画像間の違いをはっきりさせ，選択を容易にする。

提案手法を実際にシステムとして実装して三つの実験を行った。まず一つめは，ユーザの入力からユーザが違いをはっきりと感じる直線が推定できるかを確認するための実験を行った。デザイナーの感性語に関するアンケートの結果を入力に用い推定した直線と，判別分析によって得られる特徴量空間内での二つの感性語に対応する画像のクラスタを最もはっきりと分離する直線に対して，それぞれの直線上に画像を射影したときの分布と，直線上での二つのクラスタのクラスタ間分散とクラスタ内分散の比を比較した。その結果，入力によって得られた直線上の分布，分散の比ともに判別分析で得られる直線と近いものが得られ，ユーザの違いの感じ方を反映した直線が推定できることが確かめられた。

二つめの実験として，直線に沿った候補の絞りこみが有効かどうかを確認する実験を行った。提案手法に基づき実装したシステムと，提案手法と同じインタフェースを持ち，絞りこみに特徴量空間内での距離を用いるシステムの二つを用いて，二つのシステムに同じ入力を適用したときの入力回数に対する適合率と再現率の変化を調べた。その結果，本手法が従来手法より優れていることが確認できた。

三つ目の実験では，4 人の被験者に実際に提案システムと二つめの実験で用いた比較用のシステムで検索を行ってもらい，検索結果・絞りこみの分かりやすさ共に提案手法の方が良いという評価を得た。

# A Similar Image Retrieval System Using Users' Feeling of Differences between Images

Takeshi UEDA

**Abstract**

We propose a similar image retrieval system by image browsing which utilizes users' feeling of differences between images estimated from users' inputs.

As a result of the recent improvement of computers and storage units, a large amount of images get stored in computers and the usage of images becomes diversified . Owing to these backgrounds, image retrieval methods by which users can retrieve images efficiently and easily are desired.

A similar image retrieval method by image browsing is one of those image retrieval methods. An image retrieval system by image browsing performs retrieval by repetition of showing images by the system and choice of users. In this system, users can retrieve images even if they don't have clear mental image of the pictures they want, and users can use the system easily because all users have to do is to choose images from the ones shown by the system.

There are two methods in recent researches on image retrieval methods by image browsing. One method is that users choose images from representative images of clusters generated by a system in advance based on some criteria, the other method is directly comparing images chosen by users with images in a database. But these methods have two problems. One is that the way the system narrows the candidates is not easy to understand for users, the other is that users cannot choose images correctly because the images the system shows are used both for showing the result and for accepting users' inputs.

In this paper, we propose a method to narrow the candidates and an user interface for the purpose of resolving the problems shown above. We assume that users feel difference clearly between two clusters in the feature space. One cluster includes images which users want and the other includes images which users do not want. Lines linking these two clusters reflect the users' feeling of differences. In our method, the system estimates the lines which reflect users feeling of differences by two inputs from users, one is "○ " which means that the image is near to the images users want, the other is "× " which mean that

the image is far from the images users want. By narrowing the candidates along these lines, the way the system narrows the candidates becomes easy to understand for users. And the system shows images of candidates and images for accepting users' inputs separately. The system choose the images which is far away each other along the lines in the feature space in order to clarify the differences between the images. By showing these images only for accepting users' inputs, users can choose images easily.

We implemented these methods as a prototype system and conducted three experiments. First, we did an experiment in order to confirm whether the system could properly estimate users' feeling of differences between images from users' inputs. We used as inputs the result of questionnaire about the images corresponding impressive adjectives. By comparing the lines which was generated by the system with the result of the discriminant analysis which most clearly separated the two clusters corresponding to the impressive adjectives, we concluded that the system could estimate users' feeling of differences between images sufficiently.

Second, we implemented another retrieval system which had the same interface as our system and narrowed the candidates by a distance in the feature space, which had been used in recent researches. We used the same input as the first experiment for two systems and compared the change of the precision rate and the recall rate along the proceedings of retrieval. From this experiment, we concluded that our method to narrow the candidates was better than that of recent reserches.

In the third experiment, we let four subjects use two systems used in the second experiment and compared the result of retrieval and the clarity of the method to narrow the candidates. By this experiment, we confirmed that the result of retrieval in our system is better than that of the other system and the method to narrow the candidates in our system was easier to understand than that of the other system.

# ユーザの画像間の違いの感じ方を利用する類似画像検索システム

## 目次


**第 1 章　はじめに　1**

**第 2 章　ユーザの画像間の違いの感じ方を利用する画像検索手法　3**

2.1　従来研究における候補の絞りこみ手法 . . . . . 3

2.1.1　クラスタの選択による絞りこみ . . . . . 3

2.1.2　距離計算による絞りこみ . . . . . 4

2.2　ユーザの違いの感じ方を利用した候補の絞りこみ手法 . . . . . 5

2.2.1　ユーザの違いの感じ方 . . . . . 5

2.2.2　ユーザの違いの感じ方を反映する直線の推定方法 . . . . 6

2.2.3　直線を用いた候補の絞りこみ . . . . . 7

2.2.4　選択の容易なインタフェース . . . . . 7

**第 3 章　ユーザの画像間の違いの感じ方を利用する画像検索システム　8**

3.1　システムの入出力 . . . . . 8

3.2　用いる特徴量 . . . . . 9

3.2.1　HSV ヒストグラム . . . . . 9

3.2.2　フーリエパワースペクトル . . . . . 10

3.2.3　濃度差平均 . . . . . 10

3.3　アルゴリズム . . . . . 11

3.3.1　初期設定 . . . . . 12

3.3.2　ユーザの違いの感じ方を反映する直線の推定 . . . . . 13

3.3.3　検索範囲の絞りこみ . . . . . 14

3.3.4　選択用画像・候補画像の表示 . . . . . 15

**第 4 章　実験及び考察　16**

4.1　直線の推定に関する実験 . . . . . 16

4.1.1　目的 . . . . . 16

4.1.2　方法 . . . . . 16

4.1.3　結果と考察 . . . . . 17

4.2　絞りこみに直線を用いることの有効性を評価する実験 . . . . . 19

4.2.1 目的 . . . . . 19
4.2.2 方法 . . . . . 19
4.2.3 結果と考察 . . . . . 20
4.3 任意の入力に対する検索結果と絞りこみの分かりやすさに関する実験 . . . . . 22
4.3.1 目的 . . . . . 22
4.3.2 方法 . . . . . 22
4.3.3 結果と考察 . . . . . 22
**第 5 章 おわりに** **24**
**謝辞** **26**
**参考文献** **26**

# 第1章　はじめに

本研究では，画像検索技術の中で検索キーとして画像を用いる類似画像検索の手法の一つである画像ブラウジングによる検索を対象とする。そして画像ブラウジングによる検索において，ユーザの入力からユーザの画像間の違いの感じ方を推定して候補を絞りこみ検索を行う画像検索システムを提案する。

近年，計算機の処理能力の向上や大容量記憶装置の発達などにより，大量の画像データの蓄積が可能となってきており，その中から効率よく画像を取り出す画像検索技術の必要性が高まっている。また，パソコンやインターネットの普及により，個人でもWWWを通じたマルチメディア情報の発信ができるようになるなど画像データの利用場面や用途は多様化しており，画像検索においても手軽に使えるようなインタフェースやユーザのさまざまな検索要求に対応できる手法が必要となってきている。

最も初期の画像検索システムは，あらかじめデータベース中の画像に対し管理者がキーワードを付与し，ユーザが言語をキーとして与えて検索を行うものであった。しかし，画像の持つ情報は言語に比べ多様でさまざまな言葉で表現できるため，そのすべてに対応できるようなキーワードの付加は不可能である。そのため，ユーザが指定できる情報が管理者が付加したものに限定され，ユーザの意図とキーワードが一致しない場合は検索がうまくいかないという問題点があった。

現在では，検索キーとして概略画や画像そのものを用い画像特徴により検索を行う類似画像検索の研究が盛んになされている。画像をキーとして用いることにより，言語に比べて表現の自由度が上がるため，言語では表現しにくい視覚的特徴などの情報も検索に反映することができる。

類似画像検索で用いられる主な検索方法としては，ユーザが概略画を描いてそれを検索キーとする方法と，画像ブラウジングによる方法 [1] の2つがある。前者は，ユーザが欲しい画像のイメージを絵として表現しそれを検索キーとして用いる方法であり，後者は，システムが数個から10数個の選択用画像の表示し，ユーザがその中から自分の希望する画像に近いものを選択するという手順を繰り返すことによって検索を行う手法である。後者は前者と比べてあいまいなイメージからでも検索が可能で，またユーザは単に画像を選択するだけでいいので，手軽さという面でも優れていると考えられる。

画像ブラウジングによる検索には，

1. 前もってデータベース画像を何らかの基準によりクラスタリングしておいて，各クラスタを代表する画像をユーザに選択させる手法 [2, 3]
2. ユーザの選択した画像に応じて候補を絞りこむ手法 [4, 5]

の二つがあるが，従来のこれらの手法には次のような問題点がある。

- システムが検索目標の画像をどのように絞りこんでいるのかがユーザに分かりにくい
- 検索結果が次回の選択用画像を兼ねているため，表示される画像間の違いがユーザに分かりにくくなり，ユーザが自分のイメージに合う画像を選択しづらい

また， 1 の手法についてはクラスタリングの結果にユーザが縛られてしまい，ユーザの意図を反映した検索ができないという問題点もある。

そこで本研究では， 2 のユーザの選択を絞りこみに用いる方法について，上で述べた二つの問題点を改善するための候補の絞りこみ手法とインタフェースを提案する。システムはユーザの入力から，ユーザがある観点で違いを感じる特徴を特徴量空間内の直線により推定し，その直線上で候補を絞りこむ。ユーザの違いの感じ方に応じて候補を絞りこむことにより，ユーザにとってシステムの絞りこみ方が分かりやすくなることが期待できる。また，ユーザの選択する画像を検索結果とは別に表示することにより，ユーザにとって選択のしやすい画像が選択用画像に表示できると考えられる。

以下，第 2 章で従来の画像ブラウジングによる検索システムで用いられた手法について考察する。その上で，ユーザの検索意図を反映でき，ユーザに分かりやすい絞りこみ手法と選択の容易なインタフェースを提案する。第 3 章では第 2 章での提案に基づき実装した画像検索システムの詳細について述べ，第 4 章で実装したシステムを用いてユーザの違いの感じ方を反映した直線の推定，従来の絞りこみ手法との比較，そして任意の入力に対する検索の結果と絞りこみの分かりやすさの比較の三点についての実験を行った結果について報告する。最後に第 5 章でまとめと今後の課題について述べる。

# 第2章　ユーザの画像間の違いの感じ方を利用する画像検索手法

本章では，まず従来の画像ブラウジングによる画像検索システムにおける候補の絞りこみ手法について考察し，その問題点を明らかにする。その上でユーザにとって分かりやすい絞りこみ手法として，ユーザが違いを意識する特徴を特徴量空間内の直線として推定しその上で候補を絞りこむ方法と入力の容易なインタフェースを提案する。

## 2.1　従来研究における候補の絞りこみ手法

### 2.1.1　クラスタの選択による絞りこみ

クラスタリングによる絞りこみとは，データベース画像の特徴ベクトルを前もって階層的にクラスタリングしておき，各クラスタの代表画像をユーザに選択させることによって検索を進めていく手法である。

[2] では，画像をいくつかのサブブロックに分割し各サブブロックで濃淡の平均値や標準偏差などを特徴量とし特徴ベクトルを生成し，ユークリッド距離を用いた単純クラスタリングによりクラスタリングし，各クラスタの重心に最も近い画像を代表画像としていた。

また，[3] では，顔画像データベースの画像の内の一部分を顔の部品ごとに人手で分類した結果をニューラルネットに学習させ，他の全画像をそのニューラルネットに適用しクラスタを生成する。代表画像は [2] と同じ方法を用いて表示している。

[2] では，特徴量として人間の直観では分かりにくいものが使われているので，クラスタリング結果も人間にとって分かりにくくなってしまう。また [3] では，データベース画像をクラスタリングした人の画像の分類の基準と実際のユーザの基準が同じであるとは限らないので，ユーザにとって分かりにくいクラスタリングになることがある。

さらに両方に共通する問題として，クラスタの代表画像の表示の方法がある。両者ともクラスタの重心に近い画像をそのクラスタを代表するものとしているが，クラスタの重心がクラスタ内の全画像の特徴を代表できるわけではないので，ユーザには特にクラスタの境界付近の画像がどのクラスタにあるのかがはっきりしないという問題がある。

### 2.1.2　距離計算による絞りこみ

[4] では，ユーザは表示される選択用画像に対し希望する画像に近い画像を選んでどの部分に注目するかを領域として与える。システムはその領域の情報を対称行列 $\boldsymbol{A}$ で表し，選択された画像とデータベース画像との距離を

$$(\boldsymbol{i}_k - \boldsymbol{q})^T \boldsymbol{A} (\boldsymbol{i}_k - \boldsymbol{q}) \quad (k = 1, 2, \cdots)$$

として計算する。ただし，$\boldsymbol{q}$ はユーザの選択した画像の特徴ベクトル，$\boldsymbol{i}_k$ はデータベース画像の特徴ベクトル，$\boldsymbol{x}^T$ はベクトル $\boldsymbol{x}$ の転置を表す。この距離がある一定値よりも大きいものを除外していくことによって候補を絞りこむ。この手法は，距離が二次形式で計算されており，図 1 のようにユーザの入力が $\boldsymbol{q}$，$\boldsymbol{q}'$，$\boldsymbol{q}''$ と進むにつれ特徴量空間中でだ円の形で候補が絞られていく。この絞りこみ手法では，特徴量空間内であらゆる方向から絞りこまれるので利用者にとってはどのようにシステムが候補を絞りこむのか分かりずらいと考えられる。

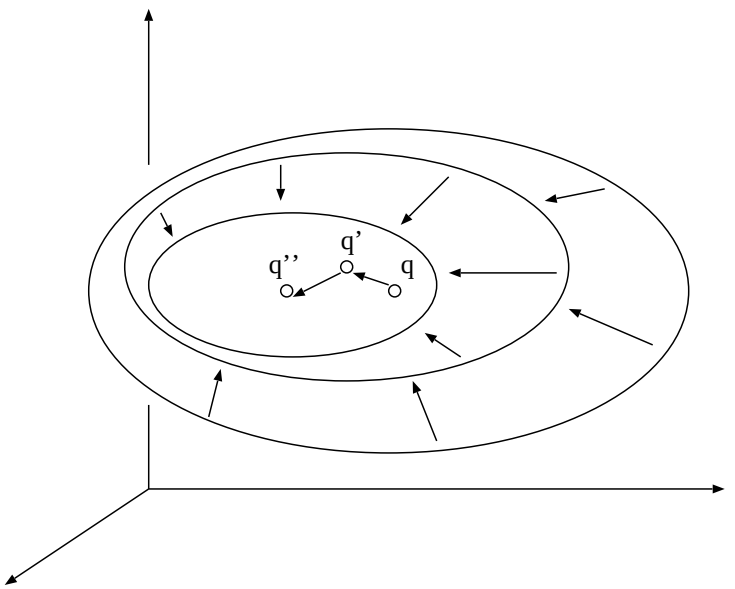



図 1: 二次形式での距離の計算による絞りこみ

また，[5] では特徴ベクトルとして画像の各ピクセルの画素値からなるベクトルを作り，主成分分析により次元を下げたものを利用する。まずユーザが選択した画像の特徴ベクトルの各要素のそろい具合である分散の値を計算する。そして，その値が小さい要素について平均値から一定範囲内にある画像に候補を絞りこんで，その上でユーザの入力した画像の平均に近い画像を検索結果とし

て表示する。この手法では，図 2 のように特徴軸に沿った絞りこみをしていると考えられるが，用いられている特徴軸は主成分分析により作られたものであり，ユーザにとって必ずしも分かりやすい特徴量ではないため，このような絞りこみ手法でもユーザにとって分かりやすいとは言えない。

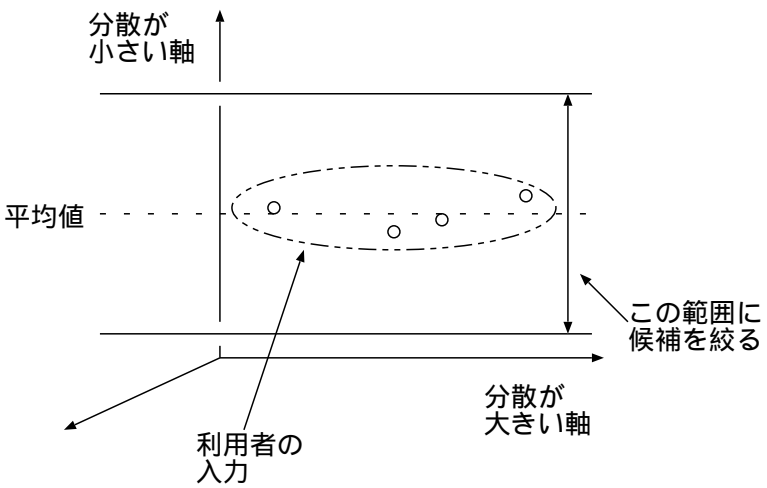



図 2: 特徴軸に沿った候補の絞りこみ

## 2.2 ユーザの違いの感じ方を利用した候補の絞りこみ手法

前節で挙げた手法は，いずれも絞りこみやクラスタリングに用いる基準，あるいは特徴量がユーザの感覚に直接結びつかないために分かりにくくなったと考えられる。また，クラスタリングによる絞りこみではユーザの選択肢がクラスタリングの結果に限定されるため，ユーザの検索意図を反映した柔軟な検索はできない。

そこで，本研究ではユーザの入力からユーザがどのような観点で違いを感じるのかを推定し，その感じ方に応じて候補の絞りこみを行う手法を提案する。この手法により，ユーザの検索意図を反映させながらユーザにとって分かりやすい候補の絞りこみが可能になる。また，検索にユーザの意図を反映することも可能になる。

### 2.2.1 ユーザの違いの感じ方

たとえば「明るい」画像には，その画像に写っている対象物が「明るい」ものや雰囲気が「明るい」ものなど，様々な観点での「明るい」画像が考えられる。

そして，それぞれの観点から見た「明るい」画像に対し反対の意味を持つような画像が存在すると考えられる。本研究ではこのような状況を想定して，特徴量空間内で図 3 で示すようにユーザは違いを感じると仮定する。

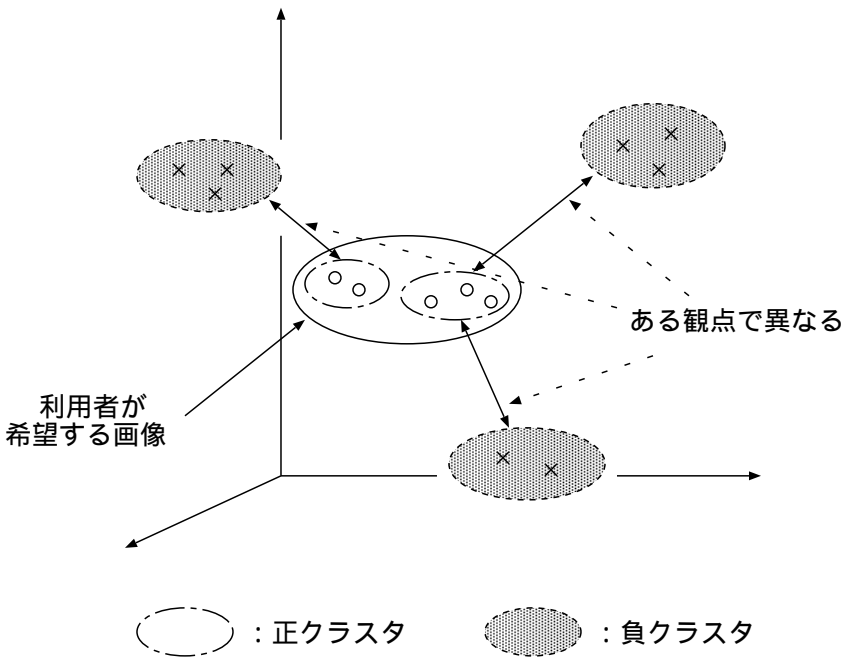



図 3: ユーザの違いの意識の仕方

ユーザが希望する画像は特徴量空間である一つのクラスタを形成している。そのクラスタ内はさらにいくつかの小クラスタ ( 本稿ではこのクラスタを正クラスタと呼ぶ ) に分かれている。そして，それぞれの正クラスタとある観点でユーザが「違う」と感じる画像のクラスタ ( このクラスタを負クラスタと呼ぶ ) が存在し，ユーザはこれらを結ぶ直線に沿って違いを顕著に感じるものと考える。

### 2.2.2 ユーザの違いの感じ方を反映する直線の推定方法

本研究で提案するシステムでは前節で述べた仮定に基づき，ユーザに希望する画像に近いものだけではなく, 希望する画像とある観点で反対の意味を持つ画像も選択させる。そして，特徴量空間内で希望する画像に近い画像と遠い画像を結ぶ直線に画像を射影したとき，分散の値が大きい直線がユーザが違いを顕著に感じる直線であるとする。分散の値が大きい直線を選ぶ理由は，分散の大きい直線上では画像が幅広く分布しておりそれだけユーザは違いを感じやすくなると考えられるからである。

### 2.2.3　直線を用いた候補の絞りこみ

ユーザが違いを感じる特徴を反映する直線上に，図 4 のように画像を射影して，正クラスタと負クラスタの間に境界点を設定する。そして，その境界点より負クラスタ側にある画像を検索対象から除くことにより絞りこみを行う。本手法はユーザが違いを顕著に感じる特徴を表す直線上で候補を絞りこむため，2.1.2 節の図 2 のような絞りこみと比べてもユーザにとって分かりやすくなることが期待できる。

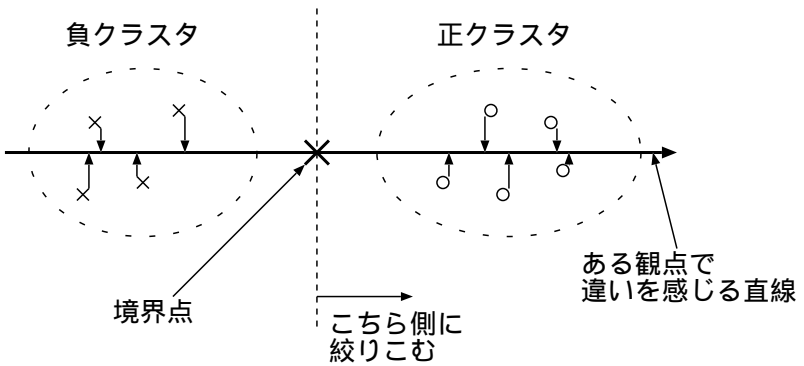



図 4: 提案手法による候補の絞りこみ

### 2.2.4　選択の容易なインタフェース

従来手法では，検索結果として表示される候補画像がユーザの選択用の画像を兼ねていた。しかし検索結果の表示において，ユーザの「似ている」という感覚とシステムの類似度の計算方法にずれが生じる可能性があるため，検索結果として表示される画像あるいはクラスタの代表画像間の違いがユーザにとって分かりにくくなるという問題がある。そのため，結果表示用の画像をそのまま選択用の画像とすると，ユーザにとって選択しづらくなってしまうと考えられる。

そこで本手法では，図 5 のような形で，2.2.2 節で推定した直線上に検索範囲内にある画像を射影したときに，直線上の分布の両端に位置する画像と中央に位置する画像を組にして表示する。ユーザが何らかの観点で違いを感じる特徴を表す直線に沿って画像を表示することにより，表示される画像間の違いが鮮明になり従来のシステムよりも選択しやすくなると考えられる。

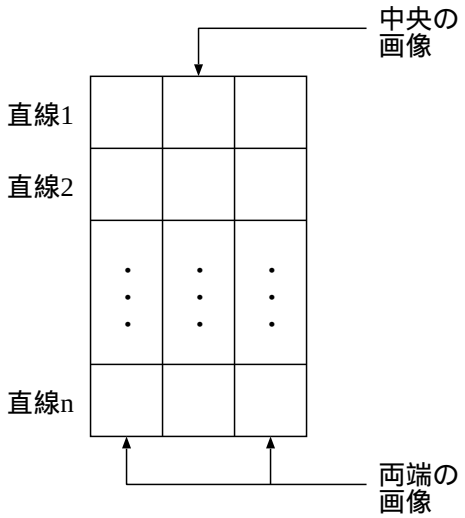



図 5: 選択用画像

# 第 3 章　ユーザの画像間の違いの感じ方を利用する画像検索システム

本章では，第 2 章で提案した手法に基づいて実装した画像検索システムの詳細を説明する。

本システムの概要を図 6 に示す。ユーザの入力を受けてシステムはユーザが違いを感じる特徴を特徴量空間内の直線として推定する。そしてユーザが実際に違いを感じていると判断した直線に沿って候補を絞りこんで，選択用画像と候補画像をユーザに提示する。この処理を繰り返し，ユーザが希望する画像が候補画像に現れた時点で検索が終了する。

## 3.1　システムの入出力

システムの入力はユーザが選択用画像につける印で，ユーザは選択用画像の中の希望する画像に近いと思うものに○ ，希望する画像とは何らかの観点で反対の意味を持つと思うものに× の印をつける。ただし，ユーザは○ ，× ともに最低一つはつけるものとする。

ユーザの入力からシステムはユーザが違いを感じる特徴を反映する直線を推定し，検索範囲の絞りこみを行い新たに選択用画像と候補画像を出力する。

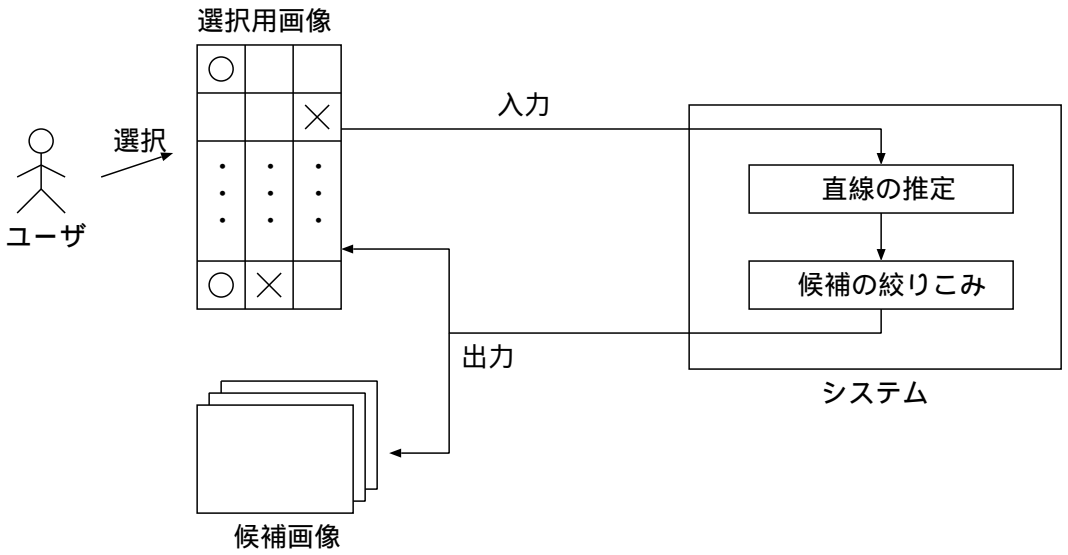



図 6: システムの概略図

## 3.2 用いる特徴量

本研究では，[6] で用いられたものと同じ以下の特徴量を使用して特徴量空間を生成する。

- HSV ヒストグラム
- フーリエパワースペクトル
- 濃度差平均

以下の各節でそれぞれの計算方法について述べる。

### 3.2.1 HSV ヒストグラム

画像の色彩的特徴を抽出するために，RGB 表色系の画素値を HSV 表色系に変換し，HSV の各画素値を 3 段階に量子化した HSV ヒストグラムを求める。

RGB 表色系から HSV 表色系への変換を以下の式により行う。

$$
\begin{aligned}
V &= \max(R,G,B) \\
S &= \frac{V-\min(R,G,B)}{V} \\
H &= \cos^{-1}\left(\frac{2R-G-B}{2\sqrt{(R-G)^2+(R-G)\cdot(G-B)}}\right) \\
&\quad \text{if } (B>G) \text{ then } H = 2\pi - H
\end{aligned}
$$

また，画像の構図の特徴を反映させるため，画像を $2\times2$ の領域に分割し，それぞれの領域ごとにヒストグラムを計算し，計 36 個の特徴量を抽出する。

なお，表色系に HSV 表色系を用いるのは，RGB 表色系よりも HSV 表色系の方が人間の感覚に近いと考えられるからである。

### 3.2.2 フーリエパワースペクトル

画像中のテクスチャの細かさと方向を反映する特徴量として，フーリエパワースペクトルを用いる。

$M \times N$ の濃淡画像 $f(x,y)(0 \leq x,y \leq M-1)$ の離散フーリエ変換 $F(u,v)$ は次の式で与えられる。

$$F(u,v) = \frac{1}{M^2} \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} f(x,y) \exp\left[-\frac{2\pi j}{M}(ux+vy)\right]$$

このとき，パワースペクトル $P(u,v)$ は，

$$P(u,v) = |F(u,v)|^2$$

で計算される。

この $P$ を $(u,v)$ 平面の極座標 $(r,\theta)$ で表し，$P(r,\theta)$ とする。極座標の変換には以下の式を用いる。

$$\begin{aligned} r &= \sqrt{u^2+v^2} \\ \theta &= \tan^{-1}\frac{u}{v} \end{aligned}$$

$r$ はテクスチャの細かさを，$\theta$ はテクスチャの方向性をそれぞれ表す。

$r, \theta$ をそれぞれ 6 段階に量子化し，各段階のパワースペクトルの総和を計算し，12 個の特徴量を求める。

### 3.2.3 濃度差平均

$M \times M$ の濃淡画像 $f(x,y)$ $(0 \leq x,y \leq M-1)$ を $n \times n$ 個の小領域に分割し，各領域ごとに画素値の平均値を求め，その値を新たに画素値とした画像 $f'(X,Y)$ を計算する $(0 \leq X,Y \leq \lfloor \frac{M}{n} \rfloor - 1)$。

$f'(X,Y)$ について，ある画素 $(X,Y)$ から $\delta = (\Delta X, \Delta Y)$ 離れた画素との濃度差

$$f'_\delta(X,Y) = |f'(X,Y) - f'(X+\Delta X, Y+\Delta Y)|$$

を計算し，画像全体について $f'_\delta(X,Y)$ のヒストグラムを求める ( 255 階調 )。$(X,Y)$ を画像全体に動かしたとき，このヒストグラムで濃度が $i$ となる確率を

$p_\delta(i)$ とすると，隣接する領域との濃度差の平均値 $MEAN$ は

$$MEAN = \sum_{i=0}^{255} ip_\delta(i)$$

で計算できる。$\delta$ については，画像全体を走査して計算を行っているので，$\delta = (1,1)$，$(1,0)$，$(1,-1)$，$(0,-1)$ の 4 方向について計算する。

この値を，$n = 1,4$ それぞれについて計算し，8 個の特徴量を得る。

以上により計算した 56 個の特徴量をそれぞれ平均 0，分散 1 に正規化した上で 56 次元の特徴ベクトルとする。

## 3.3 アルゴリズム

図 7 に検索の流れを示す。

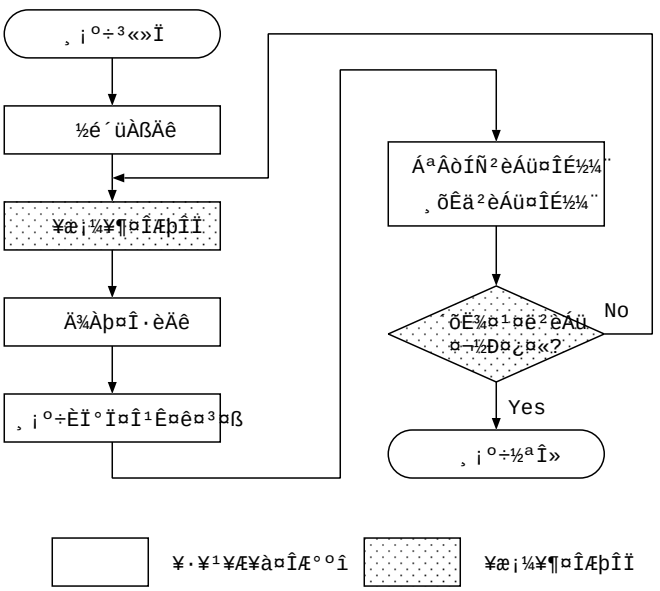



図 7: 検索の流れ

検索開始時，システムは特徴量空間の分散の大きい方向に沿って選択用画像を表示する。ユーザはシステムの表示した選択用画像に対し○ ，× の印をつける。システムはユーザのつけた印から違いを感じる特徴を反映する特徴量空間内の直線を推定する。また，システムは 1 本の直線から 3 個の選択用画像を表示するが，その 3 個の中に○ のついたものと× のついたものの両方が存在すれば，その直線に沿って検索範囲を絞りこむ。検索範囲を絞りこんだ後，システムは推定した直線上に検索範囲内にある画像を射影して選択用画像，候補画像

を決定する。

以下，

- 初期設定
- ユーザの違いの感じ方を反映する直線の推定
- 検索範囲の絞りこみ
- 選択用画像・候補画像の表示

それぞれの処理のアルゴリズムを説明する。

### 3.3.1 初期設定

本システムはユーザの○ × の入力によってユーザが違いを感じる特徴を特徴量空間内で直線として推定する。この際，初期段階においてもユーザにとって選びやすく，また絞り込みに必要な直線が早く推定できるような選択用画像が表示されるのが望ましい。

そこで， 2.2.2 節で述べたように違いの感じやすさの基準として分散を用いることとし，特徴ベクトルに対し主成分分析を適用し，特徴量空間内で分散の大きい方向を取り出す。

データベースの全画像数を $M$，各画像の特徴ベクトルを $\boldsymbol{x_i}(i = 1, 2, \cdots, M)$，全画像の平均特徴ベクトルを $\boldsymbol{\mu}$ とするとき，全画像の特徴ベクトルの分散共分散行列 $\boldsymbol{\Sigma}$ は次式で計算できる。

$$\boldsymbol{\Sigma} = \frac{1}{M-1}\sum_{i=1}^{M}(\boldsymbol{x_i} - \boldsymbol{\mu})(\boldsymbol{x_i} - \boldsymbol{\mu})^T$$

この分散共分散行列 $\boldsymbol{\Sigma}$ の固有値を大きいものから 10 個求め，それぞれに対応する固有ベクトルを計算する。この 10 個の固有ベクトルの中からランダムに 5 個を選び，原点を通る 5 本の直線を決定する。

決定した直線それぞれについて全画像を射影する。画像の直線への射影は次式で計算できる。

$$x'_{ij} = \frac{(\boldsymbol{x}_i - \boldsymbol{p})\cdot\boldsymbol{d}}{||\boldsymbol{d}||^2}d_j + p_j \quad (j = 1, 2, \cdots, n) \tag{1}$$

ここで，$\boldsymbol{x}'_i = (x'_{i1}, x'_{i2}, \cdots, x'_{in})^T$ は射影後の画像の特徴ベクトル（ $n$ は次元数 ），$\boldsymbol{d} = (d_1, d_2, \cdots, d_n)^T$ は直線の方向ベクトル，$\boldsymbol{p} = (p_1, p_2, \cdots, p_n)^T$ は直線が通っている点の位置ベクトルである。また，$\boldsymbol{x}\cdot\boldsymbol{y}$ はベクトル $\boldsymbol{x}$ と $\boldsymbol{y}$ の内積を表す。射影後の画像を特徴ベクトルの第 1 要素によりソートして 1 番目，$M/2$ 番目，

$M$ 番目になった画像 3 個を組にして，合計 15 個の画像を最初の選択用画像として表示する。ただし，15 個のうちに同じ画像が表示されないようにする。

### 3.3.2　ユーザの違いの感じ方を反映する直線の推定

- クラスタリング

  2.2.1 節の仮定に基づき，ユーザが○ をつけた画像，× をつけた画像をそれぞれクラスタリングする。クラスタリングの距離の尺度にはユークリッド距離を用い，アルゴリズムは単純クラスタリングを用いる。

  単純クラスタリングのアルゴリズムは以下の通りである。

  1. ○ のついた画像の特徴ベクトルから 1 個を選び，それをクラスタ中心とする。
  2. 他のすべての○ のついた画像について，各クラスタ中心との距離を計算する。その最小値がしきい値 $T_{circle}$ 以下であれば，最小の距離のクラスタにその画像を属させる。そうでなければ，その画像を新たなクラスタの中心とする。

  × のついた画像についてもしきい値 $T_{cross}$ を用いて同様にクラスタリングをする。

- 直線の決定

  選択用画像には，1 本の直線につき直線上の分布の両端に位置する画像と中央に位置する画像の 3 個の画像が組として表示される。その 3 個の画像の組の中にユーザが○ をつけた画像と× をつけた画像の両方が含まれる場合，システムはその直線上でユーザが違いをはっきり意識していると判断する。またこの直線を用いて 3.3.3 節で述べる絞りこみを行う。

  さらに，システムは上で述べた以外のユーザが違いを感じる特徴を反映する直線の候補として，ユーザが○ をつけた画像が属するクラスタと× をつけた画像が属するクラスタを結ぶ直線をとる。

  候補が多数ある場合は次の基準によって直線を決定する。

  1. あらかじめ各段階で○ のみが二つ以上ついた直線の方向ベクトルを履歴として保存しておく。○ のみが多数つく直線上ではユーザはあまり違いを感じないと考えられるので，この直線の方向ベクトルとの角度が小さい直線は候補から除外する。
  2. 1. で絞りこめない場合は，候補となる直線上に検索範囲内にある全画像を射影したときの特徴ベクトルの第 1 要素の分散が大きいものを順

番に選んで決定する。

以上により，最大 4 本の直線をユーザが違いを感じる特徴を反映した直線とする。なお，上の基準で候補から漏れた直線は履歴として保存しておき，直線の候補が足りない場合には履歴に残っている直線を新しいものから加えられるようにしておく。

### 3.3.3 検索範囲の絞りこみ

3.3.2 節で述べたように，1 本の直線につき直線上の分布の両端に位置する画像と中央に位置する画像の 3 個の画像が組として表示されるが，その 3 個の画像の組の中にユーザが○ をつけた画像と× をつけた画像の両方が含まれる場合，その直線を用いて絞りこみを行う。

検索範囲内にある全画像を 3.3.1 節の式 ( 1 ) により直線上に射影する。そして，× のついた画像が射影された点から直線上で距離 $dist$ 離れた点を境界として× 側に射影された画像を検索範囲から除外する。以下，$dist$ の求め方について述べる。

まず，$dist$ の最大値 $max_dist$ を決定する。射影後の○ のついた画像と× のついた画像間の距離を $t_dist$ として，ユーザが○ をつけた画像と× をつけた画像が両端のものである場合は $max_dist = K_a \cdot t_dist$，それ以外の場合では $max_dist = K_b \cdot t_dist$ とする ( $K_a, K_b$ は定数 )。

$k$ 回目の入力で，絞りこみを行う直線が $n$ 本 $(1 \leq n \leq 4)$ であるとき，$dist$ を次式で計算する。

$$dist = \left( \frac{n-1}{4} + \frac{k'}{4T} \right) \cdot max_dist$$

ただし，

$$k' = \begin{cases} k & (k \leq T \text{ のとき}) \\ T & (k > T \text{ のとき}) \end{cases}$$

ここで $n$ は $k$ 回目の選択時に決まる絞りこみに用いる直線の本数で，絞りこみに用いる直線が多いほどユーザは直線上で違いをはっきり認識できるものとして，$max_dist$ に値が近づくようにしている。また，$T$ は検索回数に関係する定数である。選択回数が多くなれば直線はよりユーザの違いの感じ方を反映できるものと考え，絞りこみに用いる直線の本数を基準にして決められた距離に加えて，選択回数に応じて $dist$ の値が大きくなるようにしている。図 8 は $n = 2$，$k' = 3$，$T = 8$ の時の $dist$ の値の決定の仕方を表したものである。

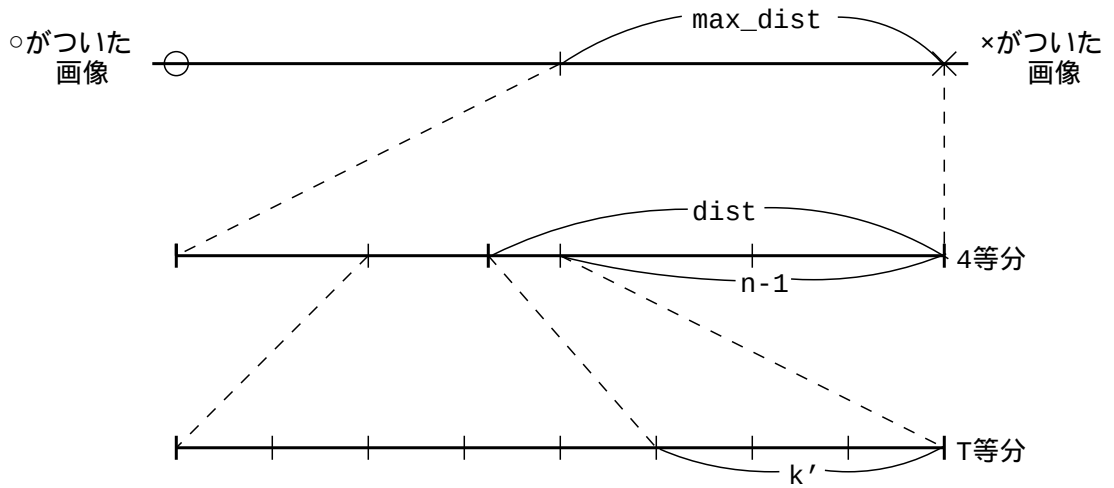



図 8: 候補を絞りこむ境界点の決定

### 3.3.4 選択用画像・候補画像の表示

- 選択用画像

  検索範囲内にある全画像を 3.3.2 節で決定した直線それぞれに 3.3.1 節の式 ( 1 ) により射影する。そして，射影後の画像の特徴ベクトルの第 1 要素で検索範囲内の画像をソートし 1 番目，$M_k/2$ 番目，$M_k$ 番目になる画像 3 個を組にして表示する。ただし，$M_k$ は $k$ 回目の選択により絞りこみが行われた後の検索範囲内にある画像数である。

  さらに，ユーザが希望する画像のイメージがあいまいな状態の時は検索意図が検索中に変わる可能性があると考えられるので，3.3.2 節で決定した直線に加えて，原点を通る直線をランダムに選んで，同様に 3 個の画像を決める。検索中に検索の進行とは無関係な画像を選択用画像に含めることにより，ユーザの検索意図の変化にも対応できると考えられる。また，3.3.1 節と同様同じ画像は表示されないようにする。

  以上により合計 15 個の画像を決定し，選択用画像とする。

- 候補画像

  3.3.2 節で決定した直線の中で，履歴から加えた直線を除いた $n$ 本の直線それぞれについて，各直線ごとにユーザが○ をつけた画像とそれに近い画像を $15/n$ 個ずつ選んで合計 15 個の候補画像を表示する。

  余りについては絞りこみに使用した直線と，直線の決定に用いた分散の値が大きかった直線から 1 個ずつ選び出すものとする。

# 第4章　実験及び考察

第 3 章で述べた検索システムを以下の 3 点について評価する実験を行った。

1. ユーザの違いの感じ方の推定について
2. 絞りこみに直線を用いることの有効性について
3. 任意の入力に対する検索結果と絞りこみの分かりやすさについて

以下，それぞれの実験について述べる。

## 4.1　直線の推定に関する実験

### 4.1.1　目的

本研究が提案するシステムでは，選択用画像として表示される画像は，検索範囲内にある画像を直線に射影したとき分布の両端に位置する画像と中央に位置する画像である。両端と中央という限られた部分しか表示に用いないため，ユーザが選択できる画像は限定されることになる。このような状況でも的確にユーザの意図を反映する直線が得られるかどうかを確かめる。

### 4.1.2　方法

まず入力として，[6] で用いられたアンケートの結果を用いる。このアンケートは，感性語による画像検索システムに意味的関係を学習させるための教師画像を得る目的で，デザイナーに市販の画像データベースから選ばれた 392 枚の風景画像から感性語にふさわしいと思う画像を選んでもらうという形で行われたものである。

感性語にふさわしいと回答のあった画像群を図 3 における一つのクラスタとみなして，一つの感性語に対応する画像群( 正クラスタ )に含まれる画像が選択用画像に現れれば○ をつけ，その感性語と違う観点を表す感性語に対応する画像群 ( 負クラスタ ) に含まれる画像が現れれば× をつける。これを入力ができなくなるまで，つまり選択用画像に○ × 両方をつけられなくなるまで繰り返す。

また，比較のために判別分析による直線の推定を行う。判別分析では，特徴量空間内で正クラスタと各負クラスタ間を最もはっきりと分ける平面に垂直な直線が求まる。この直線上に両クラスタの画像を射影したとき，クラスタ間分散/クラスタ内分散の値が最も大きくなる。

正クラスタに属する画像数を $n_1$，負クラスタに属する画像数を $n_2$ として，以下の式によりクラスタ内分散共分散行列 $\Sigma_W$，クラスタ間分散共分散行列 $\Sigma_B$ を

計算する。

$$\begin{aligned}\Sigma_W &= \sum_{k=1}^{2} \frac{n_k}{n_1+n_2}\Sigma_k \\ \Sigma_B &= \sum_{k=1}^{2} \frac{n_k}{n_1+n_2}(\overline{\boldsymbol{x}}_k - \overline{\boldsymbol{x}}_A)(\overline{\boldsymbol{x}}_k - \overline{\boldsymbol{x}}_A)^T\end{aligned}$$

ここで，$\Sigma_k$ は各クラスタに属する画像の分散共分散行列，$\overline{\boldsymbol{x}}_k$ は各クラスタに属する画像の平均ベクトル，$\overline{\boldsymbol{x}}_A$ は両クラスタに属する画像の平均ベクトルである。

このとき，二つのクラスタを最もよく分離する平面の法線ベクトルは，固有値を $\lambda$，固有ベクトルを $\boldsymbol{a}$ として，次の一般化固有値問題

$$\Sigma_B \boldsymbol{a} = \lambda \Sigma_W \boldsymbol{a}$$

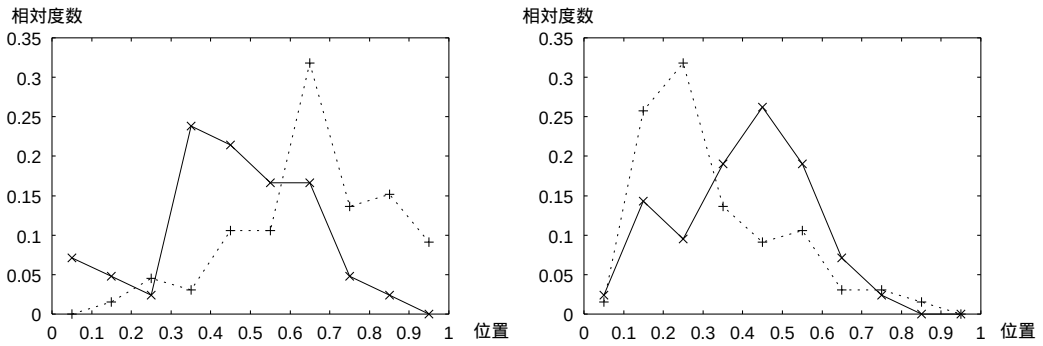
の最大の固有値に対応する固有ベクトルとなる。またこの最大の固有値がクラスタ間分散/クラスタ内分散の最大値となる。

まずデータベース画像全体を入力によって得られた直線上に射影し，分布の両端に位置する画像の特徴ベクトルの第 $1$ 要素がそれぞれ $0$，$1$ になるように全ての画像の特徴ベクトルの第 $1$ 要素を正規化する。そして，第 $1$ 要素の区間 $[0,1]$ を $10$ 等分して，それぞれの区間に正クラスタと負クラスタそれぞれに属する画像がどれだけ分布しているかを度数分布として求め，判別分析から得られる直線と比べて両クラスタに属する画像が直線上でどの程度分離できるかを調べる。また，入力によって得られた直線上に射影された画像のクラスタ間分散/クラスタ内分散の値を計算し，判別分析によって得られる値と比較する

#### 4.1.3　結果と考察

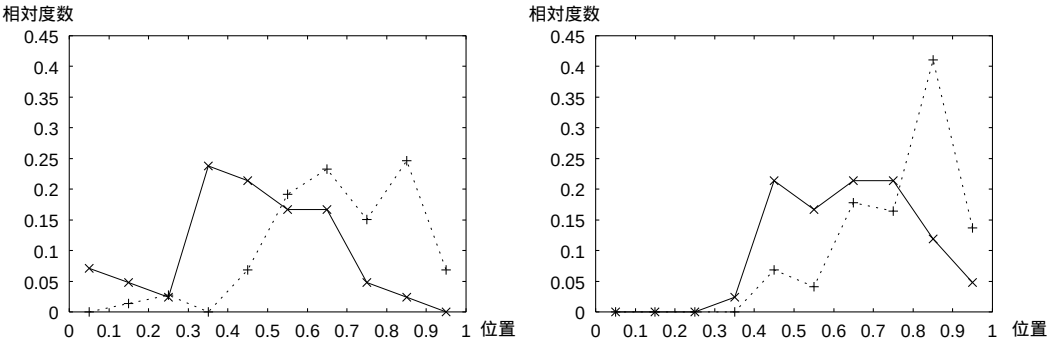
ここでは，一例として○ の入力に「 にぎやか 」という感性語に対応するクラスタ，× の入力に「 さびしい 」¦「 暗い 」という感性語に対応するクラスタを用いて上で述べた実験を行った結果を示す。$4$ 回の入力操作により，選択用画像に × の入力をするべき画像がなくなった。「 にぎやか 」のクラスタと「 さびしい 」のクラスタの組 ¦「 にぎやか 」のクラスタと「 暗い 」のクラスタの組それぞれについて上に述べた方法に従い分布を求めた結果の内，クラスタ間分散/クラスタ内分散の値が最も大きくなった直線上での分布と，判別分析で得られた直線上での分布を以下の図 $9$ および図 $10$ に示す。なお，図の横軸は直線上での画像

の位置を，縦軸は相対度数を表しており，また，実線が「にぎやか」のクラスタ，破線が「さびしい」あるいは「暗い」のクラスタの画像の分布である。




(a) 入力から得られた直線上の分布　　(b) 判別分析で得られた直線上の分布

図 9: 「にぎやか」のクラスタと「さびしい」のクラスタの画像の分布 ( 実線 :「にぎやか」の分布，破線 :「さびしい」の分布 )




(a) 入力から得られた直線上の分布　　(b) 判別分析で得られた直線上の分布

図 10: 「にぎやか」のクラスタと「暗い」のクラスタの画像の分布 ( 実線 :「にぎやか」の分布，破線 :「暗い」の分布 )

また，入力によって得られた 4 本の直線上に画像を射影したときのクラスタ間分散/クラスタ内分散の値と，判別分析によって得られたクラスタ間分散/クラスタ内分散の最大値を表 1 に示す。

直線上での画像の分布に関しては，入力によって得られた直線上の分布は，判別分析によって得られた直線上の分布と比べて多少左右に広がっているが，二

表 1: クラスタ間分散/クラスタ内分散の値

| 直線 | 「にぎやか」–「さびしい」間 | 「にぎやか」–「暗い」間 |
|---|---|---|
| 1 | 0.229001 | 0.315903 |
| 2 | 0.157225 | 0.212799 |
| 3 | 0.252027 | 0.339796 |
| 4 | 0.283885 | 0.381119 |
| 判別分析の結果 | 0.380306 | 0.446689 |

つのクラスタは同程度に分かれている。図で示した直線以外についても「にぎやか」–「さびしい」間の 2 本目の直線では重なる部分が多かったものの，ほかの直線では分離ができていた。また，クラスタ間分散/クラスタ内分散の値についても，判別分析によって得られた値に近い値が入力によって得られている。

また，この実験において入力回数は 4 回で，○ をつけた画像はのべ 10 個，× をつけた画像はのべ 13 個であった。これに対して判別分析に用いたアンケート結果のそれぞれの感性語ごとの画像数は「にぎやか」が 42 個「さびしい」が 64 個「暗い」が 73 個であり，判別分析に用いた画像数よりも少ない画像数でも，判別分析によって得られる直線に近いものが得られており，限られた入力からでもユーザの違いの感じ方を反映した直線が推定できたといえる。

## 4.2 絞りこみに直線を用いることの有効性を評価する実験

### 4.2.1 目的

本研究で提案するユーザが違いを感じる特徴を反映した直線を用いた候補の絞りこみ手法が，直線を用いない絞りこみ手法に比べて性能面で有効かどうかを評価する。

### 4.2.2 方法

第 3 章 に従って実装したシステムとの比較のため，候補の絞りこみ手法 ( 検索範囲の限定方法 ) のみが異なる検索システムを実装した。このシステムでは直線を用いない絞りこみ手法として，× のついた画像に近い画像を検索範囲から除く手法を用いる。ただし，検索範囲から除かれる画像数を提案手法を用いた場合と同数になるようにする。

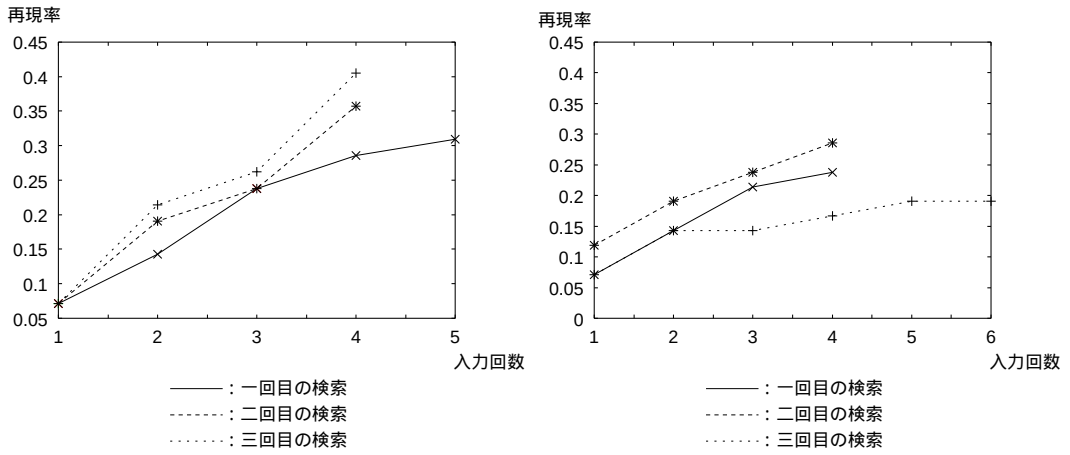
提案システムと比較用のシステムに対し 4.1 節と同様にアンケート結果を入

力し，正クラスタに対応する画像を正解画像としたときの入力回数と再現率( 出力された正解数/全正解数 ) と適合率 ( 出力された正解数/のべ出力数 ) の関係を比較する。

ただし，ここでは一度候補画像に表示された正解の画像はそれ以降候補画像には現れないようにした。このようにする理由は，候補画像の表示そのものは単に○ のついた画像に近いものを用いているため，検索が進んでも候補画像に現れる画像が変化しなくなってしまい，再現率・適合率が変化しなくなる可能性が考えられるからである。

#### 4.2.3 結果と考察

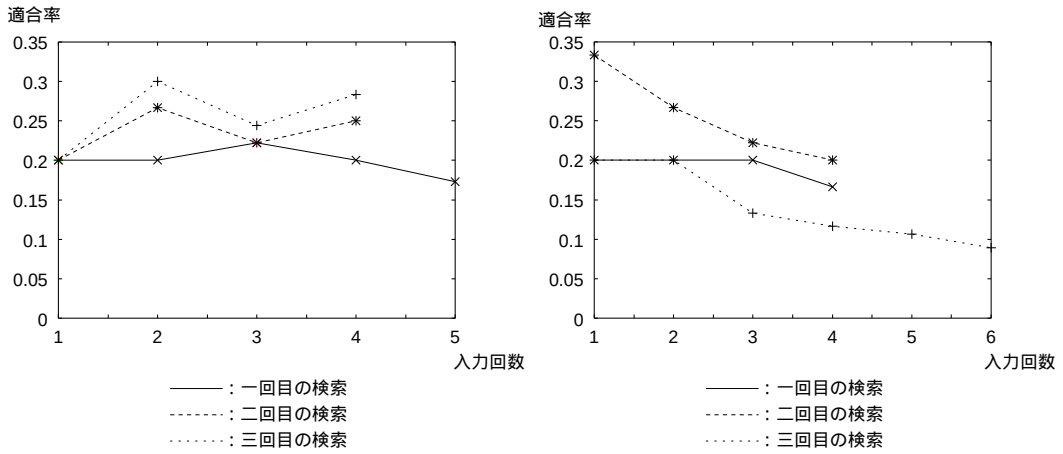
二つのシステムでそれぞれ 3 回の検索を行ったときの入力回数と再現率の関係，入力回数と適合率の関係をそれぞれ以下の図 11 および図 12 に示す。図の横軸は入力回数，縦軸は再現率あるいは適合率をそれぞれ表しており，また図中の 3 本のグラフがそれぞれ 1 回ずつの検索の結果に対応する。




(a) 提案システムによる検索結果　　(b) 比較用のシステムによる検索結果

図 11: 二つのシステムの検索時の入力回数と再現率の関係

提案システムによる検索の方は，比較用のシステムと比べて入力回数が増えるときの再現率の上昇の度合が大きく，また，適合率が低下する度合が小さくなっている。入力の回数が増えても提案システムでは正解の画像を出し続けることが可能であるのに対して，比較用のシステムでは入力回数の増加に従って表示される正解数が減少してしまう。このことから，絞りこみに直線を用いるこ




(a) 提案システムによる検索結果　　(b) 比較用のシステムによる検索結果

図 12: 二つのシステムの検索時の入力回数と適合率の関係

とにより直線を用いない絞りこみよりもよい検索結果が得られることが分かる。

しかし，両方に共通する問題として全体的に再現率・適合率がともに低いということがある。適合率が最も高い時でもその値は $0.33$ で，これは結果表示用の $15$ 個の画像の内 $5$ 個しか正解が表示されていないということになる。実際に表示される画像を見ると，正解ではないが正解のクラスタに含まれても良いような画像が含まれることが多くあった。また正解のクラスタとは全く関係のない画像も数個見られた。

今回実装したシステムでは色やテクスチャといった特徴を用いたが，同じような色やテクスチャをもっていても画像が表す意味が異なるということが起こったと考えられる。また，特徴量空間の次元が大きいため，ある画像と特徴量空間中で近くにある画像でも見た感じが大きく異なるということがあった。したがって，特徴量に画像の持つ意味を表すようなものを導入し，さらに主成分分析を用いて特徴量空間の次元を下げるなどして，より検索に有効な特徴量空間を構成する必要があると考えられる。

## 4.3 任意の入力に対する検索結果と絞りこみの分かりやすさに関する実験

### 4.3.1 目的

実際に被験者に提案システムと従来の絞りこみ手法によるシステムで検索を行ってもらい，任意の入力を用いた場合の両システムの検索結果について比較・評価を行う。また，絞りこみの分かりやすさについて比較する。

### 4.3.2 方法

まず，被験者に「にぎやか」という言葉にふさわしい画像を選んで○ をつけてもらう。そして× はその○ のついた画像と同じ直線上にある他の二つの画像から選んでつけてもらう。今回は絞りこみの様子の分かりやすさを評価するため，同じ直線上の画像に○ と× の両方をつけてもらうことにより，毎回の入力で必ず絞りこみを行うようにする。この入力の操作を，被験者が選択用画像に「にぎやか」という言葉にふさわしい画像がないと判断するまで，あるいは検索範囲中の画像数が 15 個以下になるまで繰り返してもらう。ただし，入力の操作の回数は最大で 5 回までとした。

この検索操作を二つのシステムそれぞれについて 3 回ずつ行ってもらい，そのあと検索中の候補画像の変化の様子を再び見てもらって，検索過程で出力される候補画像中に「にぎやか」という言葉にふさわしい画像が何個現れたかを回答してもらい，1 回の入力で表示される検索結果の画像中の平均の正解画像数 (のべ出力数に対する被験者が「にぎやか」という言葉にふさわしいと思う画像ののべ数の割合) を二つのシステムで比較する。また，入力のしやすさ・絞りこみの分かりやすさについての質問に答えてもらった。

### 4.3.3 結果と考察

下の表 2 に二つのシステムそれぞれについて，3 回の検索における 4 人の各被験者ごとののべ出力数に対する「にぎやか」という言葉にふさわしいと思う画像の数の割合の平均と，4 人の被験者の平均の割合を示す。

どの被験者においても提案システムでの値が比較用システムでの値を上まわっており，ユーザの任意の入力による検索においても提案する絞りこみ手法が絞りこみに直線を用いない手法と比べて良い検索結果が得られることが分かる。両システム共検索結果中に表示される正解画像の数が小さくなっているが，これは 4.2 節で述べたように特徴量空間の構成に問題があるためである。

表 2: 1 回の入力で表示される候補画像中の平均の正解画像数の割合

| | 提案システム | 比較用システム |
|---|---|---|
| 被験者 1 | 36.9% | 26.7% |
| 被験者 2 | 29.3% | 21.8% |
| 被験者 3 | 27.1% | 17.2% |
| 被験者 4 | 25.7% | 19.1% |
| 平均 | 29.8% | 21.2% |

また絞りこみの分かりやすさについては，二つのシステムそれぞれでの検索が終了した時点では，どちらが分かりやすいかをはっきり答えた被験者はいなかったが，後で選択用画像の変化の過程を見てもらった時点では，3 人の被験者が提案システムの方が分かりやすいと回答した。したがって，絞りこみの分かりやすさについても，絞りこみに直線を用いる手法の方が従来用いられていた手法よりも良いといえる。

実際に被験者に検索を行ってもらったときに，2，3 回の入力で検索範囲が絞りこまれていないにもかかわらず選択用画像に「にぎやか」にふさわしい画像が見つからなくなるということが数回あった。このようなことが，特に被験者が図 5 の真中の列の画像，つまり 1 本の直線上の 3 個の画像の内の真中の画像にのみに○ をつけたような場合に見られた。

提案システムでは選択用画像にできるだけ違いがはっきりするような画像を表示するために，直線上に検索範囲の画像を射影したときに分布の両端と中央に射影される画像を用いる。この中央の画像に○ がつきこの画像を通る直線が次回の選択用画像の表示に用いられる場合，直線の方向が変わるか，また変わらなくても絞りこみが起こって検索範囲中の画像数が変化すると，○ のついた画像の近くの画像が選択用画像に現われなくなる。また，被験者が「にぎやか」にふさわしいと思う画像が検索範囲とそれ以外の画像の境界付近にない場合にも，選択用画像に該当する画像が現れないということが起こる。

この問題を解決する方法としては，単純に直線上の分布の両端と中央にあたる画像を用いるのではなく，分布の両端付近，中央付近でそれぞれユーザが○ をつけた画像に近いものを選んで表示するといったことが必要になると考えられる。

# 第5章　おわりに

従来の画像ブラウジングによる類似画像検索手法では，システムの候補の絞りこみ方がユーザに分かりにくいということと，ユーザにとって選択しづらい画像が選択用画像に表示されるという二つの問題点があった。一つめの問題点は，システムが絞りこみに用いる基準がユーザの感覚に直接結びつかないことから生じるものであり，二つめの問題点はシステムの結果表示がそのままユーザの選択用画像になっており，表示される画像間の違いがユーザにとって分かりにくいことから生じていた。

本研究ではこの二つの問題点を改善するため，候補の絞りこみの基準として特徴量空間内でユーザの違いの感じ方を反映した直線をユーザの入力から推定し，その直線上に画像を射影して候補を絞りこむ手法と，検索結果と選択用画像を分離し，選択用画像には各画像間の違いがユーザに分かりやすい画像を表示するインタフェースを提案した。

提案に基づき検索システムを実装し，三つの実験を行った。一つめの実験では，デザイナーの感性語に関するアンケートをシステムへの入力として用いて推定した直線と，判別分析により求めた二つの感性語に対応する画像のクラスタを最もはっきりと分離するような直線に対し両クラスタの画像を射影したときの分布と直線上の分散比を比較して，ユーザの違いの感じ方を反映する直線が推定できることを確認した。二つめの実験では，提案したシステムと，従来手法で用いられた入力された画像との距離で絞りこみを行う手法を用いたシステムで入力回数に対する再現率・適合率の変化を比較し，本研究で提案した絞りこみ手法が有効であることを確かめた。さらに三つめの実験では，二つめの実験で用いた二つの検索システムを実際に使用してもらい，絞りこみの分かりやすさと任意の入力に対する検索結果の精度について比較した。そして，提案手法では従来の絞りこみ手法に比べて絞りこみが分かりやすいという評価を得た。また，検索結果についても二つめの実験と同様に提案手法の方が良くなることが確認できた。

今後の課題としては，まず従来の画像ブラウジングによる検索システムにおける候補画像と選択用画像が同一のインタフェースと本研究で提案した選択用画像を候補画像と分離するインタフェースで入力のしやすさや検索の精度を比較することがある。

また提案システムでは，クラスタ内にそのクラスタとは関係のない画像が混ざってしまうために結果として表示される画像の中にユーザの希望するものとは関係のない画像が出てしまうということがある。この問題を改善するような特徴量の取り方や特徴量空間の構成法について考慮する必要がある。

さらに，被験者に実際にシステムを利用してもらった結果，選択用画像にユーザの希望する画像に近いものが全く表示されなくなる場合があることが分かった。この問題を解決するための選択用画像の選び方の改善も今後の課題である。

## 謝辞

本研究を行うにあたり多くの御教示を賜りました池田克夫教授に深く感謝致します。そして，日頃より熱心な御指導および数多くの貴重なアドバイスを頂きました椋木雅之助手に心から感謝致します。

また，研究に際し多くの御協力・御助言を頂きました八木啓介助手，藤川賢治助手，ならびに池田研究室の皆様に深く感謝致します。

## 参考文献

[1] 坂内正夫: 画像検索技術, 電子通信情報学会誌, Vol. 71, No.9, pp. 911–914 (1988).

[2] 中嶋正臣, アンドレアクティチ, 中村太一, 富永英義: 濃淡レイアウトに基づく類似画像検索方式, 画像電子学会, Vol. 27, No.5, pp. 540–547 (1998).

[3] Wu, J.-K.: Content-Based Indexing of Multimedia Databases, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 9, No.6, pp. 978–989 (1997).

[4] Ankerst, M., Kriegel, H.-P. and Seidl, T.: A Multistep Approach for Shape Similarity Search in Image Databases, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 10, No.6, pp. 996–1004 (1998).

[5] 尾田正臣, 赤松茂, 深町映夫: あいまい顔イメージの検索に対する KL 展開の適用性, 電子通信情報論文誌, Vol. 79A, No.2, pp. 288–297 (1996).

[6] 田中大典: 対義関係にある感性語対を用いた画像検索, 京都大学工学部情報学科特別研究報告書 (1998).