

平成 28 年度
学士学位論文

色ヒストグラムと ORB 局所特徴を 組み合わせた部分画像検索

Region-Based Image Retrieval using a Combined
Feature of Color Histogram and ORB

1170320 笹谷 健文

指導教員 吉田 真一

2017 年 2 月 28 日

高知工科大学 情報学群

要 旨

色ヒストグラムと ORB 局所特徴を 組み合わせた部分画像検索

笹谷 健文

画像検索の一つとして、Region-Based Image Retrieval(RBIR) が提案されている。RBIR は、画像の部分領域を基にビジュアルキー (VK) と呼ばれる、特徴的な部分画像をクエリ情報として検索を行う手法である。RBIR の特徴表現には色特徴や局所特徴など、いくつか提案されている。しかし、これまで色特徴と局所特徴を同時に使うことは RBIR では行われてこなかった。

本研究では、色ヒストグラムと ORB 局所特徴を組み合わせた特徴表現を提案し、RBIR の性能向上を図る。色特徴に用いる色空間には、 $L^*a^*b^*$ 色空間を用いる。各部分画像の画像の色をヒストグラム化し、特徴ヒストグラムを基に部分領域をクラスタリングする。クラスタリングにより付与された各部分領域のクラスラベルを基にした類似画像の定義、および各特徴の VK を用いた検索システムの実装を行う。

被験者による評価は、200 枚の画像データセットを用いて、被験者 11 名が検索を 10 枚のテスト画像に対して行い、それぞれの適合率、再現率の調和平均である F 値で評価している。被験者実験の結果、提案手法の F 値の平均が 0.13 であり、従来手法とあまり大きな違いはないが、何人かの被験者において、提案手法が従来手法と比較して優れた値となり、それ以来の被験者ではあまり変わらない結果であり、少なくとも何人かの被験者では色特徴と局所特徴を組み合わせることが RBIR において有用であると考える。

キーワード 部分領域類似画像検索，局所特徴，大域特徴

Abstract

Region-Based Image Retrieval using a Combined Feature of Color Histogram and ORB

Takenobu SASATANI

Region-based image retrieval(RBIR) has been proposed as one of image retrieval. RBIR uses partial images as queries. Query image are called visual key(VK). Feature expressions of RBIR using various features have been proposed. However, previous studies use only one of color or local features.

In this research, we propose a combined feature of color histogram and ORB, which is a local feature in order to improve retrieval performance. For color features, $L^*a^*b^*$ color space is used. Histogram of color of pixels for a partial region of an image is computed and all histograms are classified based on the similarity of histograms. Similar images are defined based on class label of the result of clustering and VK is defined for each cluster. Proposed VK image retrieval system is implemented and evaluated.

For the evaluation of the proposed system, subjective experiment is conducted, 11 subjects performed searches on 10 test images using 200 images, and the system evaluate the F-measure as harmonic mean of the precision and recall . As a result of subjective experiment, the average of F-measure of the proposed system is 0.13, which is almost equivalent to the previous studies. However, for some subjects, the proposed method improves the F-measures and then we conclude that the proposed combined feature of color feature and local feature is useful in RBIR.

key words Region-Based Image Retrieval, Local feature, Global feature

目次

第 1 章	序論	1
第 2 章	関連研究	3
2.1	類似画像検索	3
2.2	Bag-of-Feature	3
2.3	ビジュアルキー型類似画像検索	4
2.4	従来の VK 特徴表現方法	5
2.4.1	形状特徴	5
2.4.2	色特徴	7
2.5	クラスタリング	8
2.5.1	階層的クラスタリング	9
2.5.2	非階層的クラスタリング	9
第 3 章	局所特徴と大域特徴を融合した部分類似画像検索の提案	11
3.1	部分画像の生成	12
3.2	特徴抽出	12
3.2.1	色特徴の抽出	13
3.2.2	形状特徴の抽出	14
	VW の選定	14
	VW ヒストグラムの作成	15
3.3	VW と色ヒストグラムを基にした VK の選定	15
3.3.1	色ヒストグラムと VW ヒストグラムの結合	16
3.3.2	提案手法	17
3.4	表色系が部分画像のクラスタ分布に与える影響	18

目次

3.5	部分画像データベースの構築	19
3.6	VK と適合画像の照合	19
3.7	ユーザインターフェースの設計	20
3.7.1	VK の提示	20
3.7.2	検索結果の表示	22
第 4 章	構築システムの評価	23
4.1	実験環境	23
4.2	評価指標	24
4.3	実験結果	25
4.3.1	検索精度	25
4.3.2	被験者ごとの結果	26
第 5 章	考察	29
5.1	提示する VK に関する考察	29
5.2	VW 生成に関する考察	29
5.3	部分画像クラスタリングに関する考察	30
5.4	データセットに関する考察	31
5.5	画像のサイズ変更に関する課題	32
5.6	提示した VK に関する課題	33
5.7	局所特徴のない画像に関する課題	33
5.8	次元削減に関する課題	34
5.9	データセットに関する課題	34
5.10	被験者実験に関する課題	34
第 6 章	まとめ	35
	謝辞	36

目次

参考文献

38

目次

2.1	VK 型類似画像検索	5
2.2	VK と画像データベースの照合	5
2.3	注目画素 p のコーナー検出	6
2.4	オリエンテーション α の算出	7
3.1	提案手法	11
3.2	分割画像の生成	12
3.3	RGB 値の次元削減	13
3.4	L*a*b*値の次元削減	13
3.5	VW 選定	14
3.6	VW ヒストグラムの作成	15
3.7	VW と色ヒストグラムの結合処理	16
3.8	ヒストグラムの結合での VK の選定処理	16
3.9	色の VK の選定処理	17
3.10	VW の VK の選定処理	17
3.11	提案手法での部分画像の特徴表現	18
3.12	色特徴の次元による VK のクラスタ分布	18
3.13	データベース照合	20
3.14	部分画像検索画面	21
3.15	提示される 20 個の VK(ORB 50 次元)	21
3.16	提示される 20 個の VK(L*a*b* 64 次元)	21
3.17	提示される 20 個の VK(ヒストグラム結合)	21
3.18	提示される 20 個の VK(提案手法)	22
3.19	検索結果画面	22

図目次

4.1	検索対象画像	24
4.2	各手法の適合率, 再現率, F 値 (適合率 0 を含む)	25
4.3	各手法の適合率, 再現率, F 値 (適合率 0 を除く)	26
4.4	各被験者ごとの適合率	27
4.5	各被験者ごとの再現率	27
4.6	各被験者ごとの F 値	28
5.1	類似する VK	29
5.2	VK のクラスに偏りのある特徴表現	31
5.3	Art Explosion で異なるカテゴリに属する 2 枚	31
5.4	各画像における F 値=0 が算出された回数	32
5.5	サイズ変更のない部分画像生成	32
5.6	局所特徴が少ない画像の例	33

表目次

3.1	VK データベース	19
3.2	部分画像データベース	19
4.1	実験環境	23
4.2	各手法の評価値 (適合率 0 を含む)	26
4.3	各手法の評価値 (適合率 0 を除く)	26

第 1 章

序論

インターネット上の膨大な画像データの中からユーザが要求する画像データを的確に検索することは現在でも難しい問題である．その理由として，複数ユーザの要求と検索結果における画像の印象が一致しなければならないため，画像のタグ付けを一意にするとその課題が解決できないためである．

主な画像検索手法は，メタデータを用いたテキスト検索と画像をクエリデータとして類似画像を検索する類似画像検索である．テキスト検索は，一般的に利用されている画像検索で，ユーザは要求する画像のキーワードをクエリ情報として検索する手法である．クエリ情報と画像データベース画像に付与されたメタデータを用いて照合が行われるため，メタデータが適切に付与されることが必要である．メタデータの表現は多種多様であるため，表現の仕方によって検索結果が大きく変わるという問題が生じる．

一方，類似画像検索はクエリ情報として画像を用いて，クエリ画像と特徴の類似している画像を検索する手法である．この手法はメタデータを用いないため，表現によって検索結果が変わる問題がない．しかし，ユーザがクエリ情報であるキーワード情報や類似クエリ画像を持たない場合，検索することが不可能になる．

この問題を解決するためにビジュアルキー型類似画像検索が提案されている [1]．ビジュアルキー型画像検索は，ビジュアルキーと呼ばれるあらかじめ定義されているクエリ情報を用いて，ユーザの頭の中のイメージを表現して画像検索を行う手法である．ユーザに必要なものはイメージという曖昧な情報のみで良いため，検索不可能な状況に陥りにくいのが特徴である．従来のビジュアルキーの表現として，各画像の色特徴やエッジ特徴，近年では物体認識に有効な局所特徴を用いた画像の特徴が用いられてきた．しかし，人間が物体や画像を

認識する場合，1 つの特徴で認識するのではなく色・形状・質感・大きさなど複数の要素を組み合わせるため，従来の表現手法は人間の自然な感覚とは少し異なると考えられる．ビジュアルキーは画像の部分領域の一つであり，データベースの全画像を複数に分割し，各部分画像の特徴の類似性を基に部分集合のクラスタを作成し，その中心に近い部分画像を代表画像としてビジュアルキーに選出する．本研究は，ビジュアルキー型類似画像検索をより人間の感覚に適応させることを目的とし，色特徴と局所特徴を組み合わせた特徴表現手法を提案し，より高精度なビジュアルキー型類似画像検索の構築を目指す．

提案する特徴表現手法は，色特徴と局所特徴を抽出している．色特徴は $L^*a^*b^*$ 色ヒストグラムを用いており，特徴の次元数を削減するために減色を行う．一方，局所特徴には ORB 特徴を用いており，次元を削減するために Bag of Features で表現している．単純に抽出した色特徴ヒストグラムと局所特徴ヒストグラムを結合し，1 つのヒストグラムを再構成したものを 1 枚の画像の特徴とする手法が考えられるが，局所特徴は色特徴に対して，各部分画像での抽出数が大きく異なるため，局所特徴ヒストグラムと色ヒストグラムでの尺度が異なるという問題がある．そのため，提案手法では 1 枚の画像に色ヒストグラムと局所特徴ヒストグラムの 2 つを画像の特徴として，検索時に各特徴ごとに照合する．

提案手法を用いたビジュアルキー型画像検索システムを構築し，被験者実験によって得られた適合率，再現率，F 値によって検索精度の評価を行う．被験者実験には，画像データ ArtExplosion の 10 カテゴリから 20 枚ずつ選択した 200 枚の画像を用いており，各カテゴリから 1 枚ずつ抽出した 10 枚の画像を検索対象画像としてビジュアルキーで検索を行う．F 値は， $L^*a^*b^*$ 色ヒストグラムを用いたシステムが 0.091，ORB 局所特徴ヒストグラムを用いたシステムが 0.103，ヒストグラム結合手法を用いたシステムが 0.092，提案手法が 0.127 であった．

本論文の構成は，第 2 章で本研究に関連する技術，手法について説明し，第 3 章では本研究で提案する特徴表現手法と構築システムについて説明する．第 4 章では被験者実験による構築システムの評価について述べ，第 5 章で本研究の考察および今後の課題について述べる．最後に第 6 章で本研究について簡単にまとめる．

第 2 章

関連研究

本章では、類似画像検索、およびビジュアルキー型類似画像検索に用いられた特徴抽出手法・クラスタリング手法・画像データセットについて説明する。

2.1 類似画像検索

類似画像検索は、画像の色、質感、形状などを基にして行う検索手法である。ユーザは画像をクエリとして、そのクエリ画像の特徴と類似した特徴を持つ画像が結果として算出される手法である。多くの特徴は多次元ベクトルで表現され、その特徴ベクトルを基にクラスタリングを行い、複数のクラスタを形成する。しかし類似画像検索において、視覚的データから抽出される情報とそのデータを見たときのユーザの解釈が異なるというセマンティックギャップ問題がある [2]。

2.2 Bag-of-Feature

類似画像検索に用いる特徴表現方法の一つに、Bag-of-Feature がある。これは、1枚の画像を 1 つの多次元特徴ベクトルで表現する方法で、Scale-Invariant Feature Transform(SIFT), Speeded Up Robust Features(SURF), Oriented FAST and Rotated BRIEF(ORB) などの局所特徴を k 個にクラスタリングを行い、ヒストグラムを作成することで k 次元特徴ベクトルであるビジュアルワード (VW) を生成する。各部分画像の特徴表現手法を以下に示す。

2.3 ビジュアルキー型類似画像検索

1. 局所特徴を対象画像から抽出する．
2. 抽出した局所特徴を k 個にクラスタリング．
3. k 個のクラスタのセントロイドを VW として選定する．
4. 画像中の各局所特徴について距離が一番近い VW に投票することで，全画像の局所特徴を VW の次元とするヒストグラムに変換する．

2.3 ビジュアルキー型類似画像検索

ビジュアルキー型類似画像検索 (VKIR) とは，画像の一部分の要素に注目し，抽出した局所特徴量を基にクラスタリングを行う類似画像検索の一種である．従来の類似画像検索 (CBIR) では，画像全体の特徴を用いて検索しており，一つの画像に対して全体の特徴をあらかじめ知る必要があった，例えば，手元にある画像に類似したものを探すような用途には適しているが，部分的な手がかりしかないような場合は検索が難しい．そこで，VKIR では，画像の一部分のみの特徴を用い，部分的な類似性からの検索を可能にする．まず，データベース中の画像を分割して得られた部分領域画像集合から，各クラスを代表する画像を選出する．この代表画像をビジュアルキー (VK) と呼ぶ．ユーザは，VK を組み合わせてクエリを生成する．クエリで索引付けされている画像をデータベースから算出し，結果として提示する．図 2.1 に VK 型画像検索の処理フローを図 2.2 に VK と画像の照合方法を示す．図 2.1 では，まずユーザは，サーバから提示されたクエリ画像群からイメージの部分領域に近似している VK を選択する．サーバは受け取ったクエリ情報と索引情報データベースを照合して，適合した画像を画像データベースから参照し，結果を表示する．

2.4 従来の VK 特徴表現方法

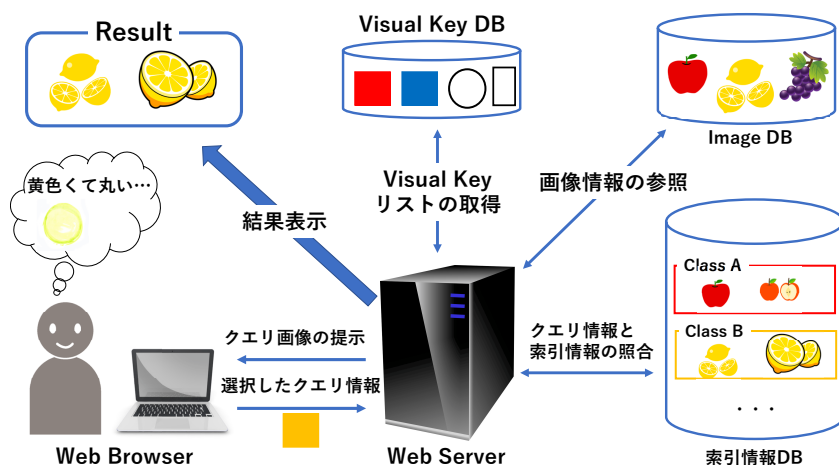


図 2.1 VK 型類似画像検索

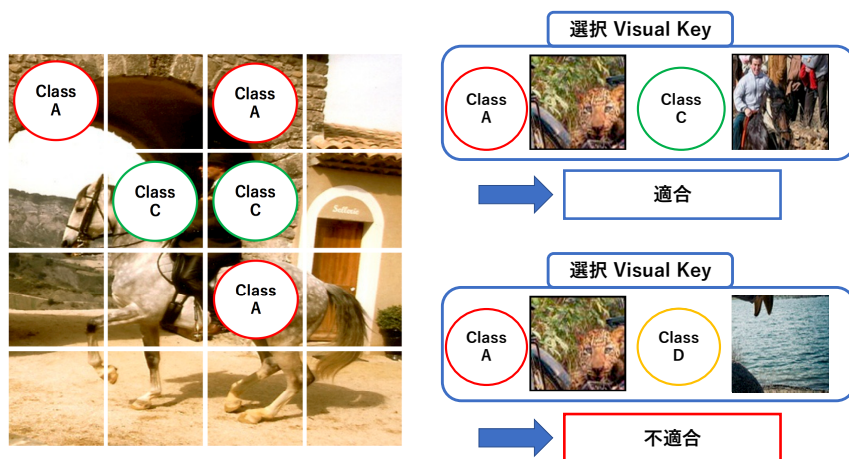


図 2.2 VK と画像データベースの照合

2.4 従来の VK 特徴表現方法

部分画像の特徴に形状特徴および色特徴を用いた VK 型類似画像検索システムがそれぞれ提案されている．それらで用いられている特徴について説明する．

2.4.1 形状特徴

画像の形状特徴の抽出には，Canny 法や Sobel 法を用いたフィルタを用いたエッジ特徴と SIFT，SURF，ORB などの局所特徴を用いた手法が提案されている [3][4][5]．

2.4 従来の VK 特徴表現方法

本研究で用いる ORB(Oriented FAST Rotated BRIEF) 特徴のアルゴリズムは以下の通りである [8][9] .

FAST Detector

注目画素 p と円周上の 16 画素の輝度値を比較し , 連続する n 個以上の輝度値が p よりも閾値 t 以上に明るい (もしくは暗い) 場合に p をコーナーとして検出する (図 2.3) . 検出速度の向上のため , 学習用画像を用いて周囲 16 画素を ' 明るい ' , ' 類似 ' , ' 暗い ' の 3 値に分類し , 決定木を作成する . 決定木に従って周囲の画素値を順に検査することで , コーナー (キーポイント) かそうでないかを識別する . この手法によって , 全画素値をチェックする必要がないので SIFT や SURF より高速なキーポイント検出ができる .

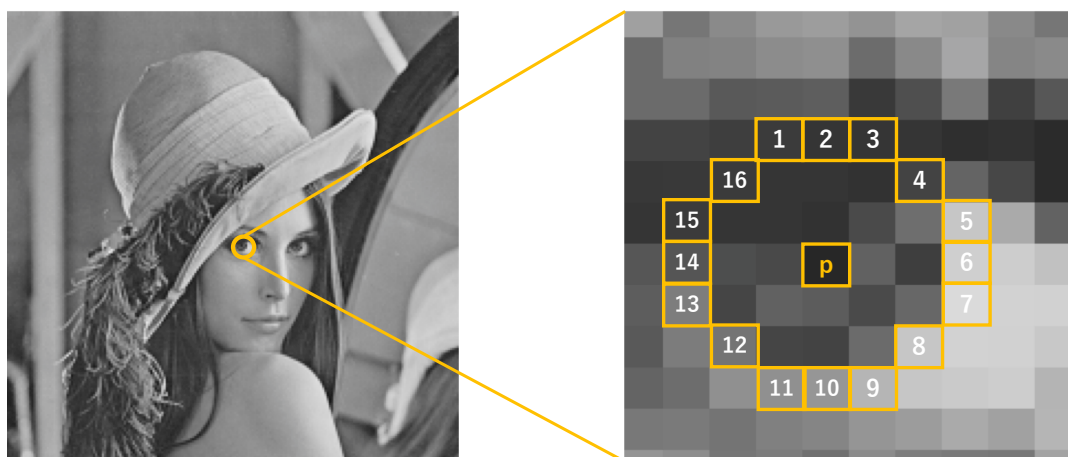


図 2.3 注目画素 p のコーナー検出

Orientation by Intensity Centroid

検出したキーポイントとその周囲 16 画素内のパッチ領域の画素値から , 0 , 1 次モーメントからパッチ領域の重心 C を算出する , キーポイント中心 p と算出した重心 C を結ぶ直線の傾きをオリエンテーション α として用いる (2.4) . 特徴量記述の際に , 輝度差を求める 2 点のペアの座標位置を α だけ回転させることで , 回転不変性を持つ特徴となる .

2.4 従来の VK 特徴表現方法

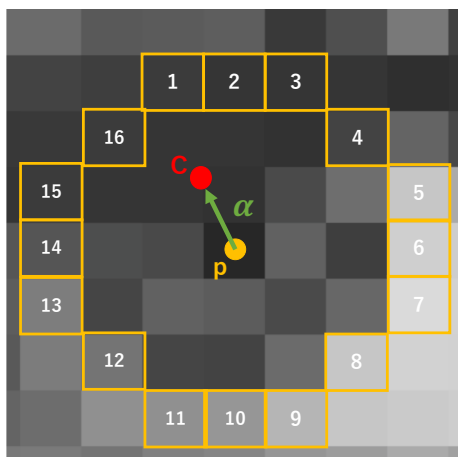


図 2.4 オリエンテーション α の算出

Efficient Rotation of the BRIEF Operator

ORB では輝度差を求める 2 点のペアを学習済みモデルにより選択する．ペアの選定基準として 2 つを定義している．

1. ビット分散が大きくなるようなペア．
2. 選択されたペア同士の相関が低いペア．

パッチ内で有り得るペアを全て列挙し，ビット分散が大きく，ペア同士の相関が低いペアを貪欲法アルゴリズムによって選択して， N 組のペアに絞りこむことで記述能力の高いバイナリコードとなる．選ばれた 2 点 p_i, q_i の輝度差 $I(p_i) - I(q_i)$ が正であれば 1，負であれば 0 を割りあてることで N ビットのバイナリコードを生成する．

2.4.2 色特徴

RGB

Red, Green, Blue の 3 つの原色を混合して単色を生成する加法混色の 1 種である．人間の網膜には 3 種類の錐体 (cone) があり，各錐体が異なる波長に選択的に反応することを用いて，国際照明委員会 (CIE) によって Red(波長:700.0nm), Green(546.1nm), Blue(435.8nm) の 3 原色が標準化された．コンピュータビジョンでは各色の値域は 0 から 255 であり， $256^3 = 16777216$ 通りの色の表現が可能である [10]．

2.5 クラスタリング

L*a*b*

約 1%の輝度の相対的な差を知覚できる人間の視覚系の反応に近くように CIE が定義した表色系である。L*(輝度), a*(色度), b*(色度) で表現される。L*の値域は $0 \leq L^* \leq 100$ であるが, a*, b*の値域に関しては変換元の表色系によって異なる, 今回用いた, RGB から L*a*b*の変換では $-127 \leq a^* \leq 127, -127 \leq b^* \leq 127$ の値域を取る。RGB から L*a*b*への変換は以下の通りである [10][11]。

注目画素の XYZ 空間を

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

で定義する。L*, a*, b*の各値は

$$L^* = \begin{cases} 116 * Y^{1/3} - 16 & Y > 0.008856 \\ 903.3 * Y & Y \leq 0.008856 \end{cases}$$

$$a^* = 500 \left(f \left(\frac{X}{X_n} \right) \right) - (f(Y))$$

$$b^* = 200 (f(Y)) - \left(f \left(\frac{Z}{Z_n} \right) \right)$$

である。このとき以下の値が用いられる。

$$f(t) = \begin{cases} t^{1/3} & t > 0.008856 \\ 7.787t + 16/116 & t \leq 0.008856 \end{cases}$$

$$X_n = 0.950456$$

$$Z_n = 1.088754$$

2.5 クラスタリング

教師なし学習法に属するクラスタリングは, データから隠れた構造を見つけ出す手法である。クラスタリングには, 大別して, 非階層的な手法と階層的な手法がある。

2.5 クラスタリング

2.5.1 階層的クラスタリング

階層的クラスタリングは、クラスタリングされていない各データを個々のクラスタとして、クラスタ間の距離が近い順にクラスタ融合をして、最終的に 1 つのクラスタに統合する手法である。融合の過程をデンドログラムと呼ばれる木の形で表現することができる。従来研究では、階層的クラスタリングの代表的手法である Ward 法が用いられている.[4]。

Ward 法

Ward 法は、クラスタ A と B の距離を融合前後のクラスタ内の各データと平均の距離の増加量と定義し、距離の小さなクラスタから融合していく手法である。Ward 法はクラスタ内変動の増加分で距離を定義しているため、階層法の中で精度の高い方法である。クラスタ内変動の増加分は以下で定義される。

$$D(A, B) = \sum_{x \in A, B} d(x, \mu_{AB})^2 - \left(\sum_{x \in A} d(x, \mu_A)^2 + \sum_{x \in B} d(x, \mu_B)^2 \right)$$

$d(x, y)$ はユークリッド距離、 μ_{AB} はクラスタ A と B を融合したクラスタの平均ベクトル、 μ_A と μ_B はクラスタ A と B それぞれの平均ベクトルである。

2.5.2 非階層的クラスタリング

非階層的クラスタリングは、階層的クラスタリングと異なりあらかじめ決まった N 個のクラスタ中心の座標を求めてそのクラスタに包含されるか否かで分類する手法である。本研究では K-means 法を用いる。

K-means 法 (K-平均法)

d 次元の N 個のデータ $D = x_1, \dots, x_N, x_i \in R^d$ をデータ間の類似性として、初期設定した K 個のクラスタに分類する手法である。K-means 法のアルゴリズムは以下の通りである。

初期化

2.5 クラスタリング

N 個のデータをランダムに K 個のクラスタに振り分け , それぞれのクラスタの平均ベクトル $\mu_k (k = 1, \dots, K)$ を求める .

帰属度変数に関する最適化

μ_k を固定し , 帰属度変数 q_{ik} を以下の式に従って決める .

$$q_{ik} = \begin{cases} 1 & k = \arg \min_j \|x_i - \mu_j\|^2 \\ 0 & \text{otherwise} \end{cases}$$

μ_k の最適化

q_{ik} を固定し , 以下の式に従ってセントロイド μ_k を求める .

$$\frac{\partial J(q_{ik}, \mu_k)}{\partial \mu_k} = 2 \sum_{i=1}^N q_{ik} (x_i - \mu_k) = 0$$

上記の式では帰属変数 q_{ik} と μ_k に関する最適化が含まれている . μ_k に関する最適化は , 以下の通りである .

$$\mu_k = \frac{\sum_{i=1}^N q_{ik} x_i}{\sum_{i=1}^N q_{ik}}$$

繰り返し

上記の”帰属度変数に関する最適化”と” μ_k の最適化”をクラスタの状態変化がなくなるまで繰り返す .

第 3 章

局所特徴と大域特徴を融合した部分類似画像検索の提案

本章では、 $L^*a^*b^*$ 色ヒストグラム特徴と ORB による VW ヒストグラム特徴を提案し、提案する特徴表現手法を用いた部分類似画像検索システムのデータベース設計、実装システム、実験方法について説明する。

提案する特徴は、生成した色ヒストグラムと VW ヒストグラムそれぞれの特徴を元に各部分画像をクラスタリングし、各部分画像が 2 つの特徴で表現される手法である (図 3.1)。

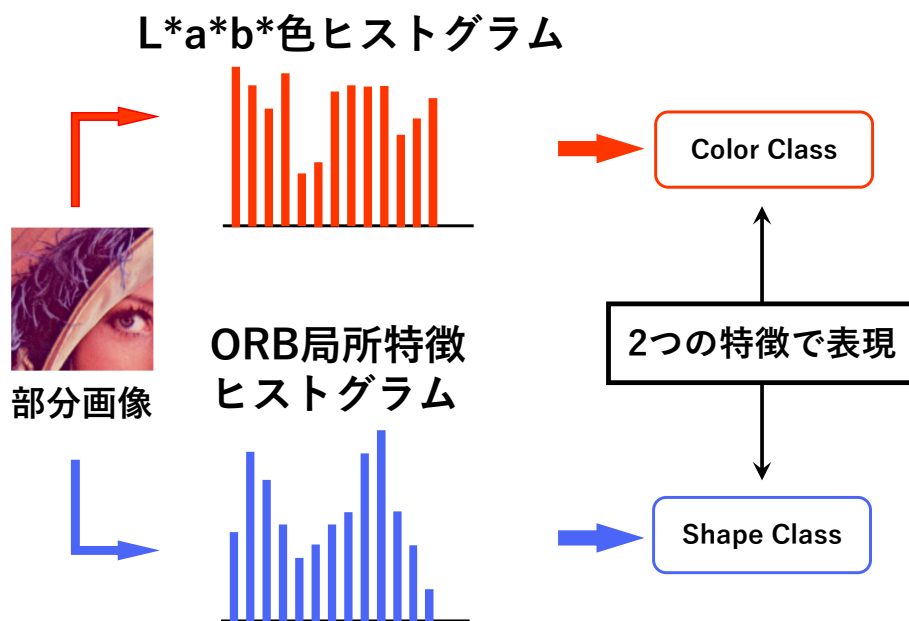


図 3.1 提案手法

3.1 部分画像の生成

システム設計の概要は以下の手順である．

1. 全画像を 480×480 画素に変換し， 4×4 に分割する．
2. 分割された部分画像に対して，RGB ヒストグラム， $L^*a^*b^*$ ヒストグラムを作成する．
3. 分割された部分画像に対して，ORB 局所特徴を抽出し，K-means 法を用いて VW ヒストグラムを作成する．
4. 作成した特徴を基に，全部分画像を K-means 法でクラスタリングする．
5. 生成したクラスタから VK を選定する．
6. データベースを構築する．

3.1 部分画像の生成

初めに，対象画像を均等に分割するために，全画像を 480×480 画素に変換する．その次に 4×4 に画像を等分割する (図 3.2) ．

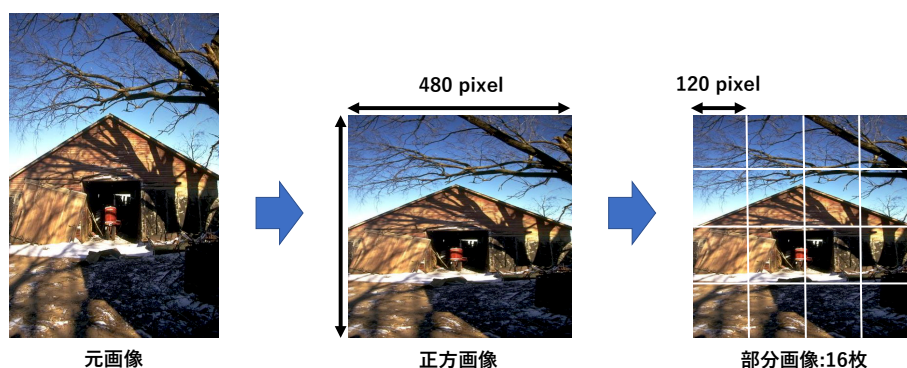


図 3.2 分割画像の生成

3.2 特徴抽出

本研究では，色特徴として RGB ヒストグラム， $L^*a^*b^*$ ヒストグラムを形状特徴として ORB 局所特徴の VW ヒストグラムを用いるため，それぞれの生成方法を述べる．

3.2 特徴抽出

3.2.1 色特徴の抽出

生成した各部分画像に対して、1つの色ヒストグラムを生成する．RGBで表現できるパターン数は、 $256^3 = 16777216$ 通りと特徴の次元が非常に大きくなるため、減色による特徴の次元削減を行う．各部分画像のRGB値を取得し、R値、G値、B値をそれぞれ n 等分し、 $0 \sim (n-1)$ を割り当てる(図3.3)．この表現方法によって $n=4$ の場合、R値、G値、B値をそれぞれ2bitで表現することができ、1枚の部分画像の色特徴は $2^{2+2+2} = 2^6 = 64$ 次元で表現することができる[6]． $L^*a^*b^*$ 色ヒストグラムの作成に対しても、RGBと同様に各値を n 等分して $0 \sim (n-1)$ を割り当てる(図3.4)．

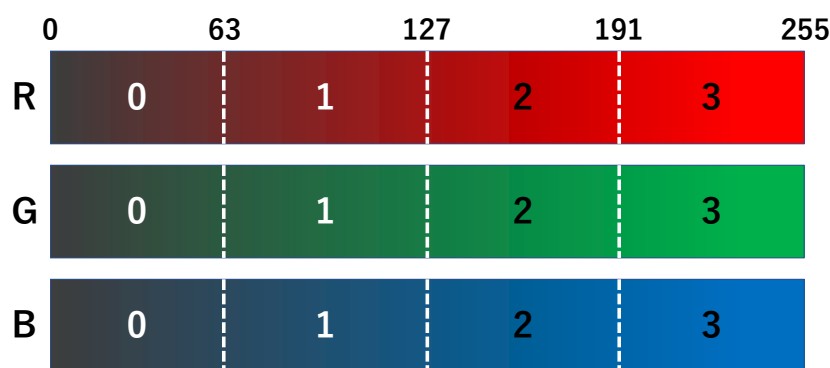


図 3.3 RGB 値の次元削減

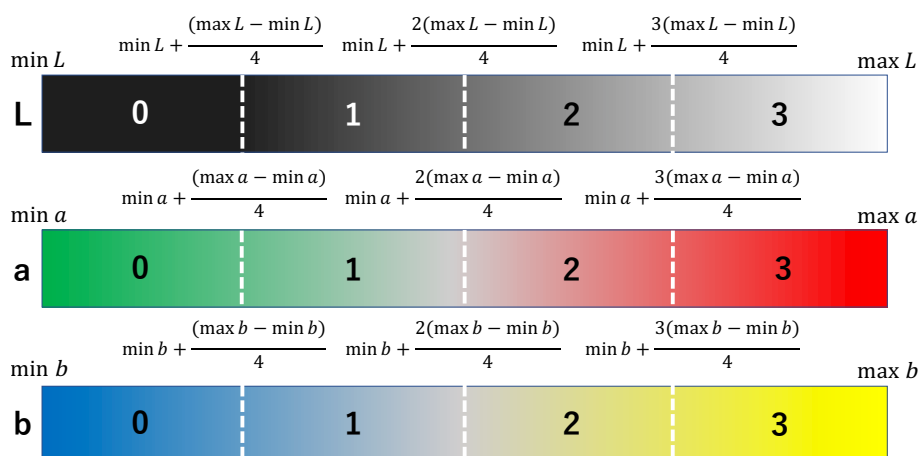


図 3.4 $L^*a^*b^*$ 値の次元削減

3.2 特徴抽出

3.2.2 形状特徴の抽出

色特徴と同様に，生成した各部分画像に対して，抽出した局所特徴から 1 つの VW ヒストグラムを生成する．VW ヒストグラムは以下の手順で生成する．

1. 抽出された全ての局所特徴を K-means 法でクラスタリングし，局所特徴の代表 (VW) を選定する．
2. 各部分画像の VW ヒストグラムを生成する．

VW の選定

VW 選定の流れを図 3.5 に示す．全部分画像から抽出された局所特徴量を K-means 法で N 個のクラスタを生成する．この時，特徴量が抽出できなかった部分画像は除外する．次に，分類された各クラスタの中心に最も距離の近い特徴を VW として選定する．

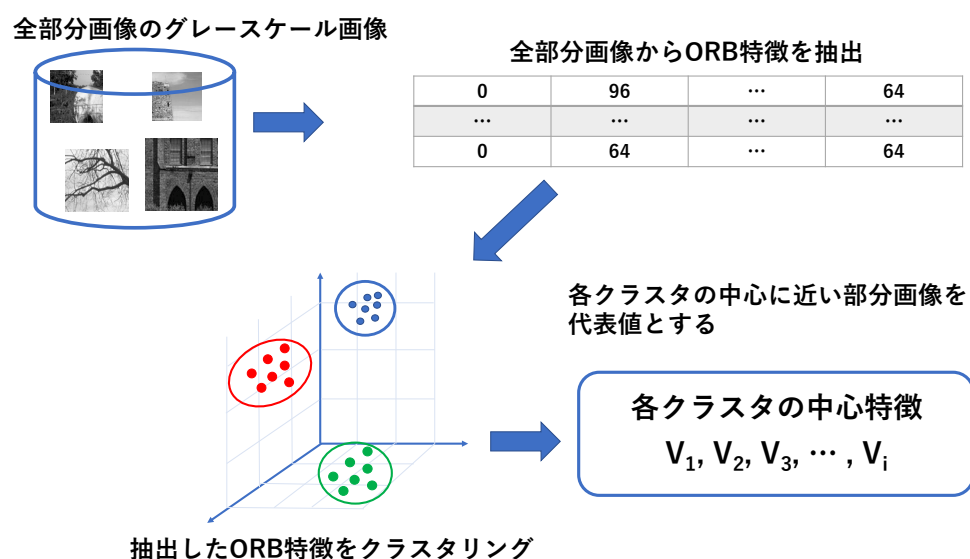


図 3.5 VW 選定

3.3 VW と色ヒストグラムを基にした VK の選定

VW ヒストグラムの作成

VW ヒストグラムの作成を図 3.6 に示す．各部分画像から抽出された各局所特徴量が属するクラスターの VW に投票することでヒストグラムを作成する．VW の数はヒストグラムの次元に相当するため，各画像の特徴を N 次元のベクトルとして表現する．

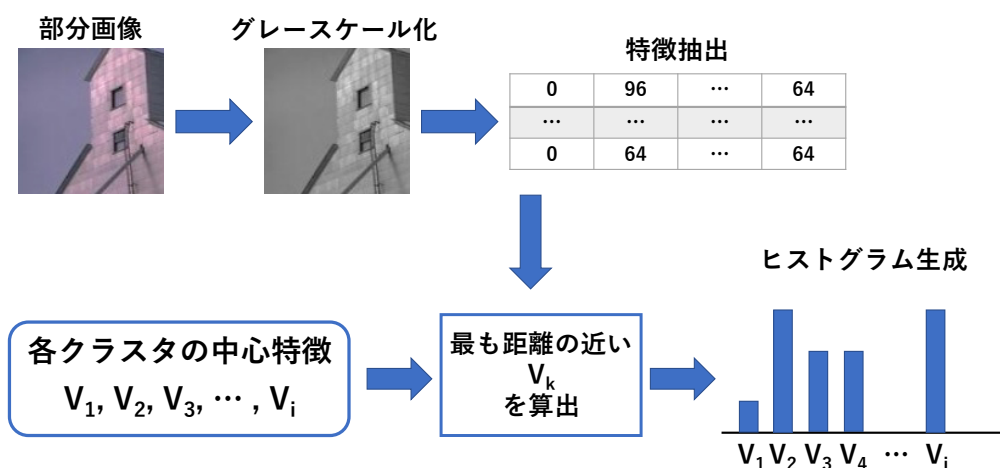


図 3.6 VW ヒストグラムの作成

3.3 VW と色ヒストグラムを基にした VK の選定

本節では，抽出した色ヒストグラムと VW ヒストグラムを基に，各部分画像をクラスタリングし，ラベル付けを行う手法について説明する．本研究では以下の 2 つの手法で両特徴を組み合わせる．

1. 色ヒストグラムと VW ヒストグラムを結合して新たな特徴表現をし，各部分画像に 1 つのラベルを付与する手法 (結合手法)
2. 各部分画像に色ヒストグラムと VW ヒストグラムのラベルをそれぞれ付与する手法 (提案手法)

両手法のクラスタリングには K-means 法を用いる．

3.3 VW と色ヒストグラムを基にした VK の選定

3.3.1 色ヒストグラムと VW ヒストグラムの結合

各部分画像の色ヒストグラムと VW ヒストグラムを結合して、新たな特徴ヒストグラムを生成し、それを基に各部分画像のクラスタリングを行いラベル付けを行う。クラスタ数に 20 を用いる。また、各クラスタの中心に最も距離の近い部分画像を VK として選定する。部分画像における色ヒストグラムと VW ヒストグラムの結合処理を図 3.7、VK 選定処理を図 3.8 に示す。

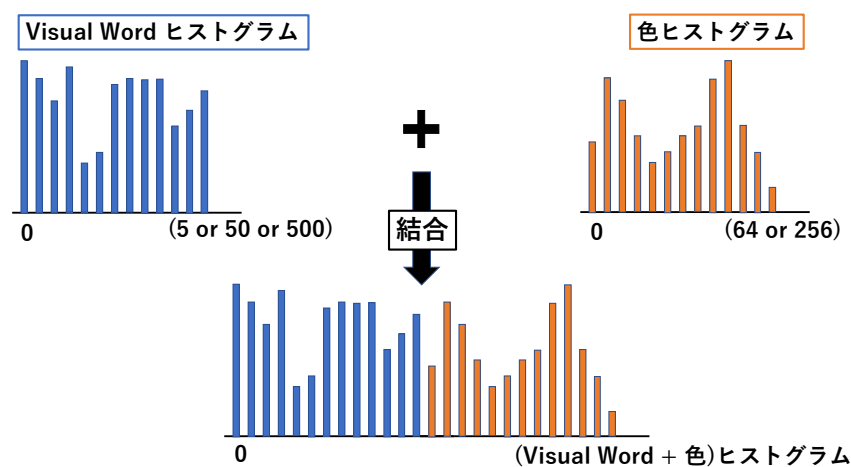


図 3.7 VW と色ヒストグラムの結合処理

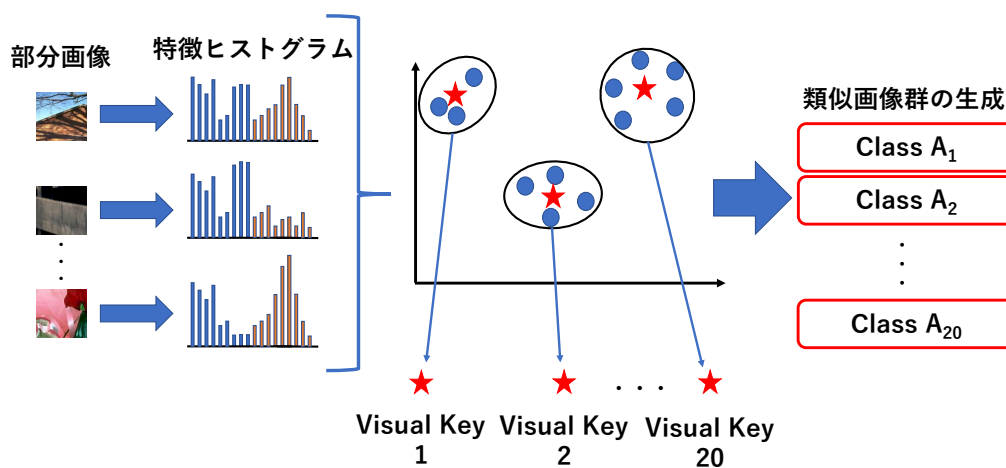


図 3.8 ヒストグラムの結合での VK の選定処理

3.3 VW と色ヒストグラムを基にした VK の選定

3.3.2 提案手法

各部分画像の色ヒストグラムと VW ヒストグラムをそれぞれにクラスタリングを行い、色と VW の 2 つのラベルを付ける．従来手法と VK の数を等しくするため、各特徴のクラスタリングに用いるクラス数は 10 とする．また、各クラスタの中心に最も距離の近い部分画像を VK として選定する．部分画像における VK 選定処理を図 3.9、3.10、3.11 に示す．

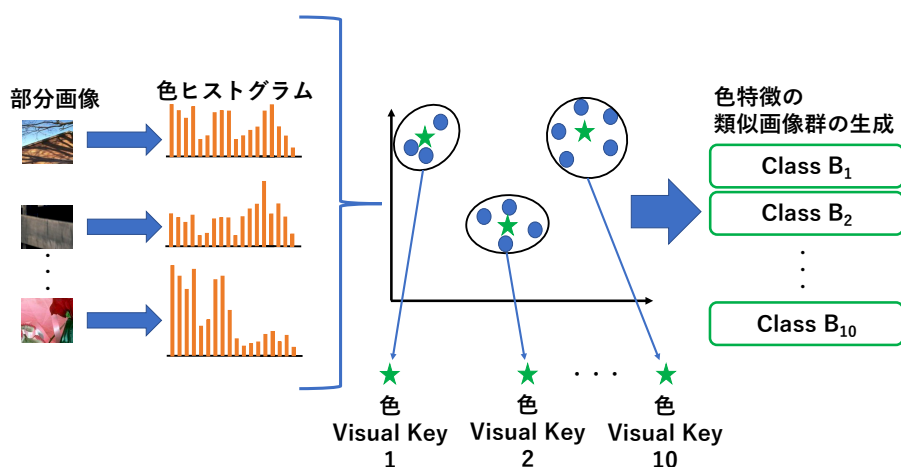


図 3.9 色の VK の選定処理

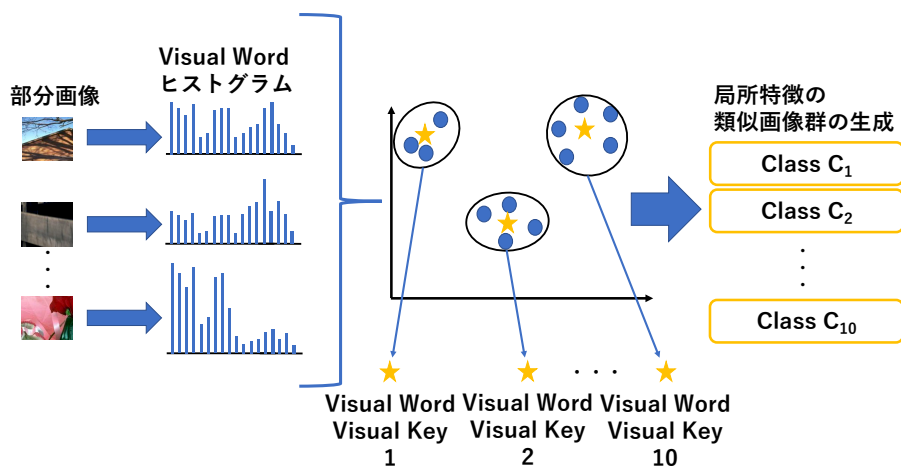


図 3.10 VW の VK の選定処理

3.4 表色系が部分画像のクラスタ分布に与える影響

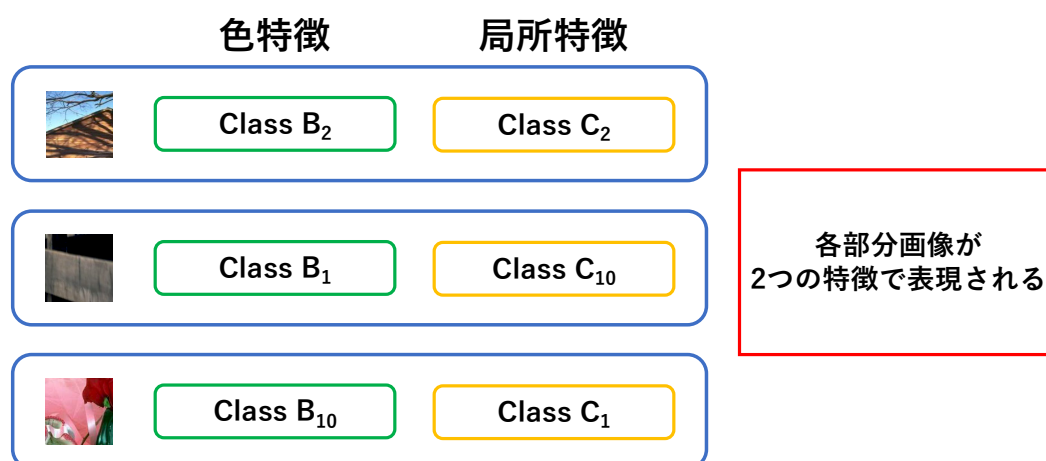


図 3.11 提案手法での部分画像の特徴表現

3.4 表色系が部分画像のクラスタ分布に与える影響

3.2 節で述べたように、本研究で用いる表色系は RGB と $L^*a^*b^*$ を 64 次元または 256 次元で減色した色特徴である。しかし、図 3.12 で示すように、色特徴を 256 次元にした場合 1 つのクラスタが巨大化してしまうため良いクラスタリング結果とならない。さらに西本の研究 [6] によって RGB より $L^*a^*b^*$ の方が F 値が向上するという結果が出されている。そのため、画像検索システムに用いる表色系は 64 次元の $L^*a^*b^*$ ヒストグラムとする。

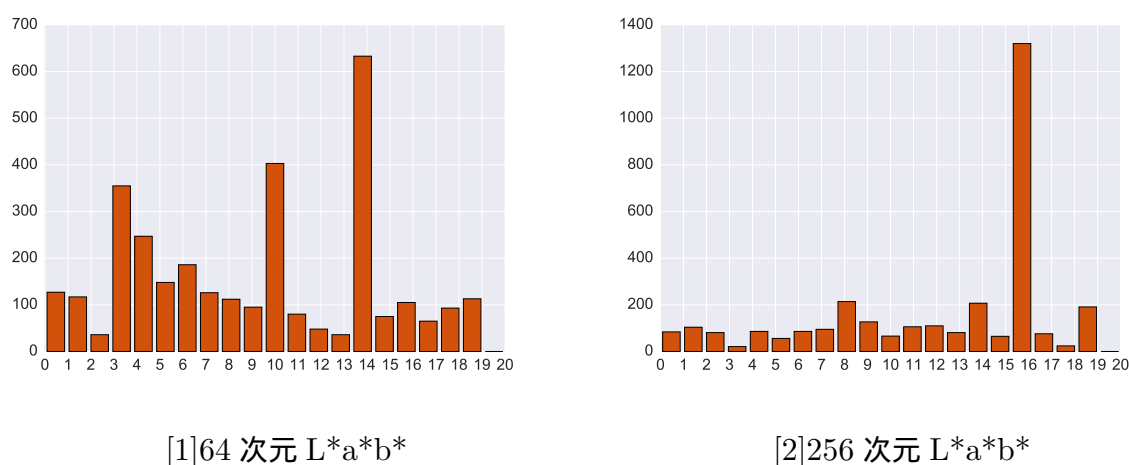


図 3.12 色特徴の次元による VK のクラスタ分布

3.5 部分画像データベースの構築

クラスタリングによって定義した，部分画像の特徴に関するラベルを基にデータベースを作成する．データベースは，各特徴表現の VK を格納する VK データベースと，各部分画像が持つラベルを格納する部分画像データベースの 2 つを構築する．VK データベースは，各特徴表現に VK を 20 個保持するために VK の部分画像 ID とクラス番号を格納している (表 3.1)．部分画像データベースは，画像検索を行う際に，VK と画像の部分領域に索引付ける役割を持ち，元画像 ID，各部分領域が持つクラスタのラベルを格納している (図 3.2)．

表 3.1 VK データベース

vk id	class
162-9	0
147-9	1
⋮	⋮
130-7	19

表 3.2 部分画像データベース

id	vk 1	⋯	vk 16
1	4	⋯	11
2	12	⋯	9
⋮	⋮	⋮	⋮
200	10	⋯	16

3.6 VK と適合画像の照合

ユーザが選択した VK から適合画像を選択する照合方法について図 3.13 に示す．ユーザが選択した VK と同じクラスの部分画像を持つ画像を適合画像として選択する．例として，ユーザがクラス 2 と 8 の VK を選択した場合，id:1 の画像はクラス 2 とクラス 8 の両方の部分画像を含んでいるため適合画像となり，id:2 の画像はクラス 8 の部分画像を含んでいないため不適合となる．

3.7 ユーザインターフェースの設計

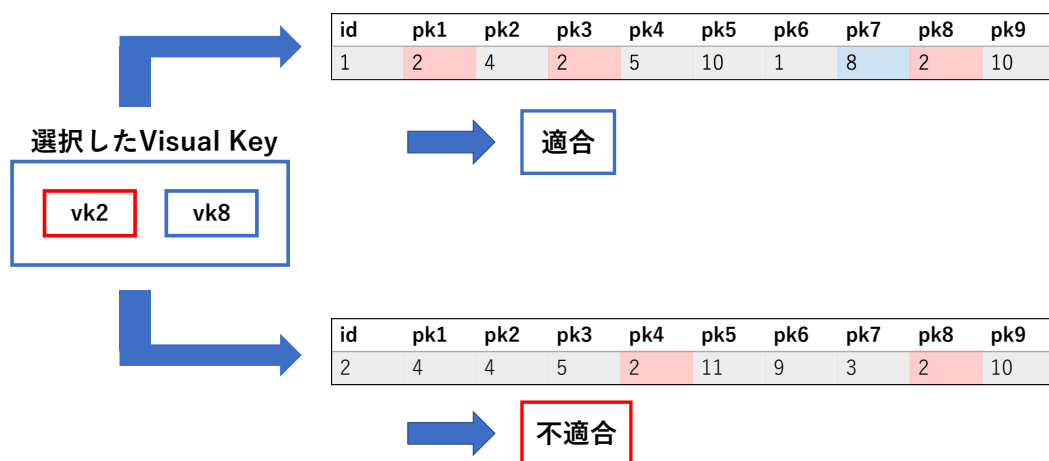


図 3.13 データベース照合

3.7 ユーザインターフェースの設計

作成したデータベースを用いて、部分画像検索システムを構築する。提案システムでは HTML および PHP を用いる。

3.7.1 VK の提示

ユーザは自分のイメージに基づいて画面下部に提示されている 20 個の VK を選択し、search ボタンで検索を行う (図 3.14)。被験者実験では、画面上部にテスト画像が表示され、ユーザはイメージに基づいて、テスト画像の検索を行う。各特徴表現の VK は図 3.15 から図 3.18 に示す。ORB、 $L^*a^*b^*$ 、ヒストグラム結合手法は 20 個の VK から選択する。提案手法は上段が ORB、下段が $L^*a^*b^*$ を表現しているため各段 10 個ずつで VK を選択する。

3.7 ユーザインターフェースの設計

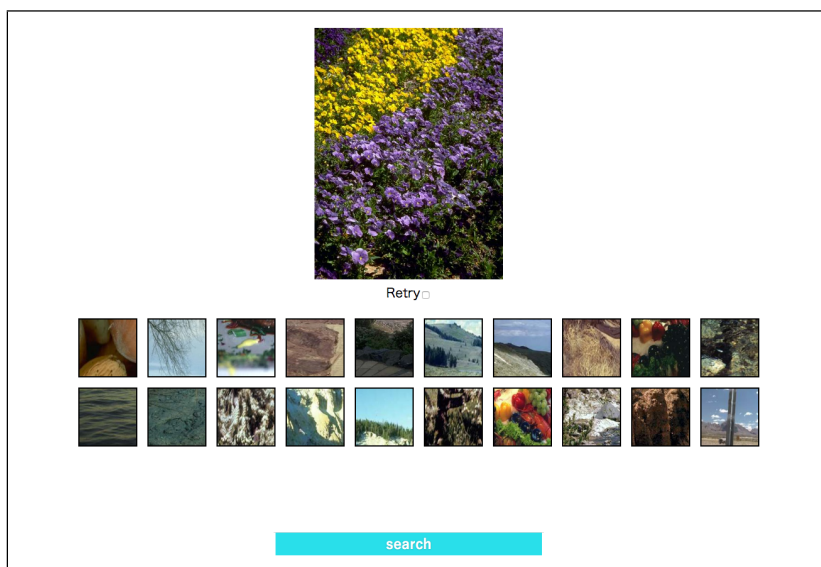


図 3.14 部分画像検索画面



図 3.15 提示される 20 個の VK(ORB 50 次元)

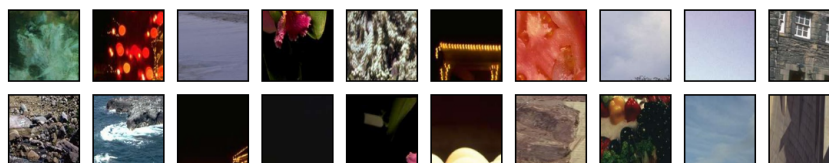


図 3.16 提示される 20 個の VK(L*a*b* 64 次元)

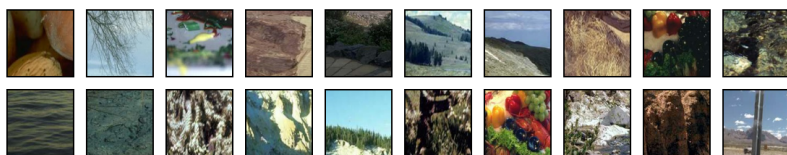


図 3.17 提示される 20 個の VK(ヒストグラム結合)

3.7 ユーザインターフェースの設計

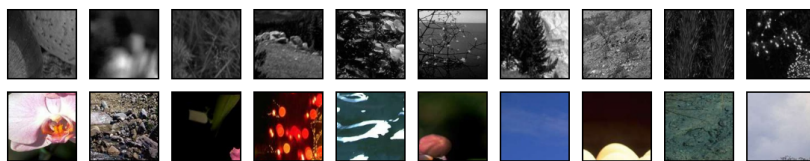


図 3.18 提示される 20 個の VK(提案手法)

3.7.2 検索結果の表示

VK を選択した後，search ボタンによってクエリ情報を Web サーバに送信する．Web サーバが受け取った VK クエリと部分画像データベースを照合することで適合画像を取得し，結果を出力する．図 3.19 は画像を検索した時に表示される結果画面の例である．

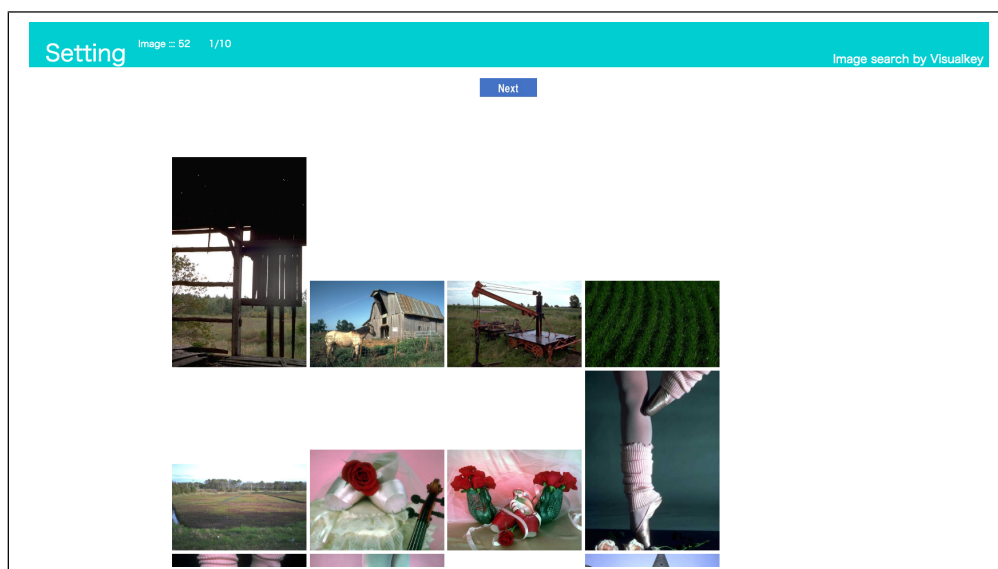


図 3.19 検索結果画面

第 4 章

構築システムの評価

本章では，従来手法である ORB VW ヒストグラム， $L^*a^*b^*$ 色ヒストグラム，両特徴の結合特徴で特徴表現した場合と提案手法で特徴表現した場合の性能を被験者実験によって行う．

4.1 実験環境

表 4.1 に実験条件を示す．被験者実験によって構築したシステムの評価を行う．被験者 11 名に図 4.1 に示す検索対象画像を 10 枚を提示し，各画像を 20 個の VK を用いて検索を試行する．選択する VK の数は 1 個以上と設定する．比較する特徴表現手法は以下の 4 パターンである．

1. ORB VW ヒストグラム (50 次元)
2. $L^*a^*b^*$ 色ヒストグラム (64 次元)
3. ヒストグラムの結合 (ORB VW ヒストグラム 50 次元 + $L^*a^*b^*$ ヒストグラム 64 次元)
4. 提案手法 (ORB VW ヒストグラム 50 次元 + $L^*a^*b^*$ ヒストグラム 64 次元)

表 4.1 実験環境

被験者	11 名
画像データ	Art Explosion 中の 10 カテゴリから 20 枚ずつ選定した 200 枚
検索対象画像	10 枚
VK の数	20 枚

4.2 評価指標



図 4.1 検索対象画像

4.2 評価指標

構築システムの性能評価として，各被験者の各条件に対して適合率，再現率，F 値を算出する．各値の算出には以下の式を用いる．

$$\text{適合率} = \frac{\text{検索結果内の適合画像の枚数}}{\text{検索結果の画像枚数}}$$

$$\text{再現率} = \frac{\text{検索結果内の適合画像の枚数}}{\text{全画像中の適合画像枚数}}$$

$$F \text{ 値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

適合率は，検索システムの正確性に関連する指標であり，再現率はシステムの網羅性に関する指標である．適合率と再現率には片方を上げればもう一方が下がるトレードオフの関係があるため，両者の調和平均を取った F 値も評価指標として用いる．また，式中の“検索結果内の適合画像の枚数”は，被験者に検索結果画像の中から検索対象画像と類似していると判断した枚数であり，“全画像中の適合画像枚数”は，被験者に全画像データから検索対象画像と類似していると判断した枚数である．

4.3 実験結果

4.3 実験結果

4.3.1 検索精度

被験者 11 人での各手法における，適合率，再現率，F 値を示す．被験者実験において，検索結果に適合画像が存在しない場合には，一度だけ再検索を行ってもらう．それでも適合画像が存在しない場合は，適合率を 0 と判定する．その結果の各評価値の平均を図 4.2 に示し，適合率 0 の結果を除いた各評価値の平均を図 4.3 に示す．

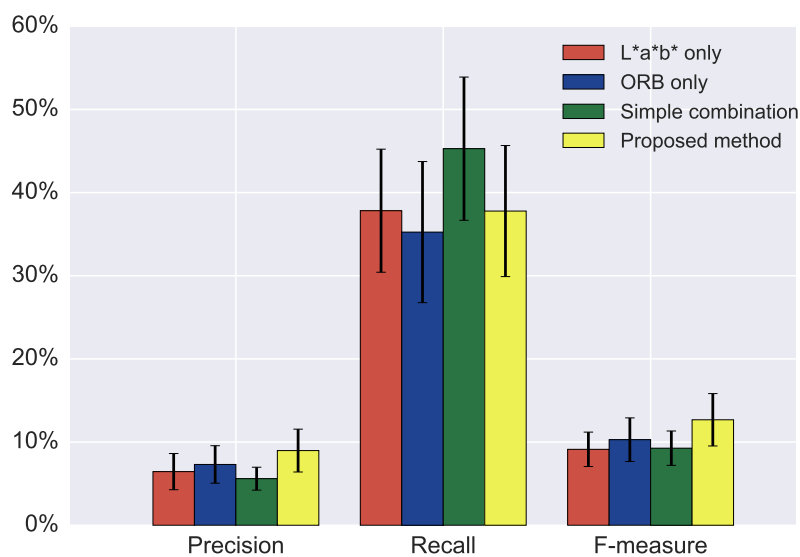


図 4.2 各手法の適合率，再現率，F 値（適合率 0 を含む）

4.3 実験結果

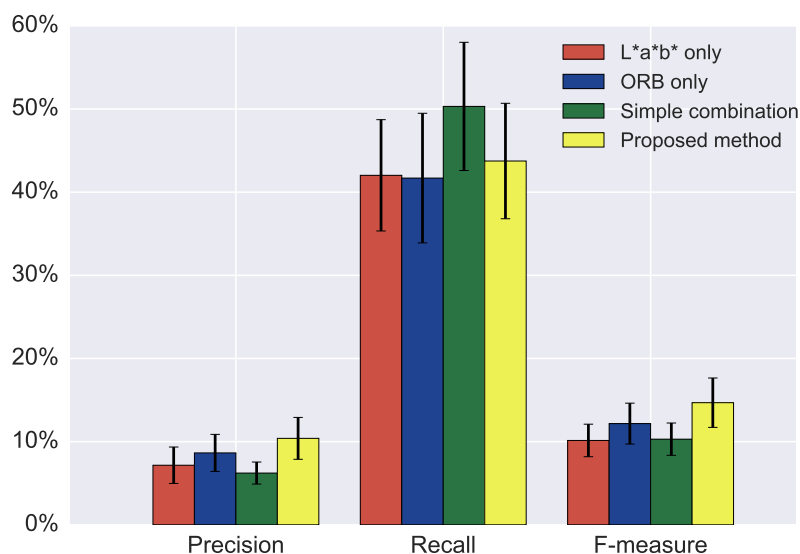


図 4.3 各手法の適合率，再現率，F 値（適合率 0 を除く）

表 4.2 各手法の評価値（適合率 0 を含む）

	適合率	再現率	F 値
L*a*b*	0.06	0.38	0.09
ORB	0.07	0.35	0.10
結合手法	0.06	0.45	0.09
提案手法	0.09	0.37	0.13

表 4.3 各手法の評価値（適合率 0 を除く）

	適合率	再現率	F 値
L*a*b*	0.07	0.42	0.10
ORB	0.09	0.42	0.12
結合手法	0.06	0.50	0.10
提案手法	0.10	0.44	0.15

4.3.2 被験者ごとの結果

検索対象画像ごとの評価を図 4.4，4.5，4.6 に表示する．図中の横軸は検索対象画像の番号，縦軸は各評価値である．

6 名の被験者において，適合率，F 値は提案手法が他の手法と同等であり，5 名の被験者は他の手法より高い値となった．適合率，F 値が低い被験者はどの手法においても低く，VK 型画像検索が不得意である傾向が見られた．

4.3 実験結果

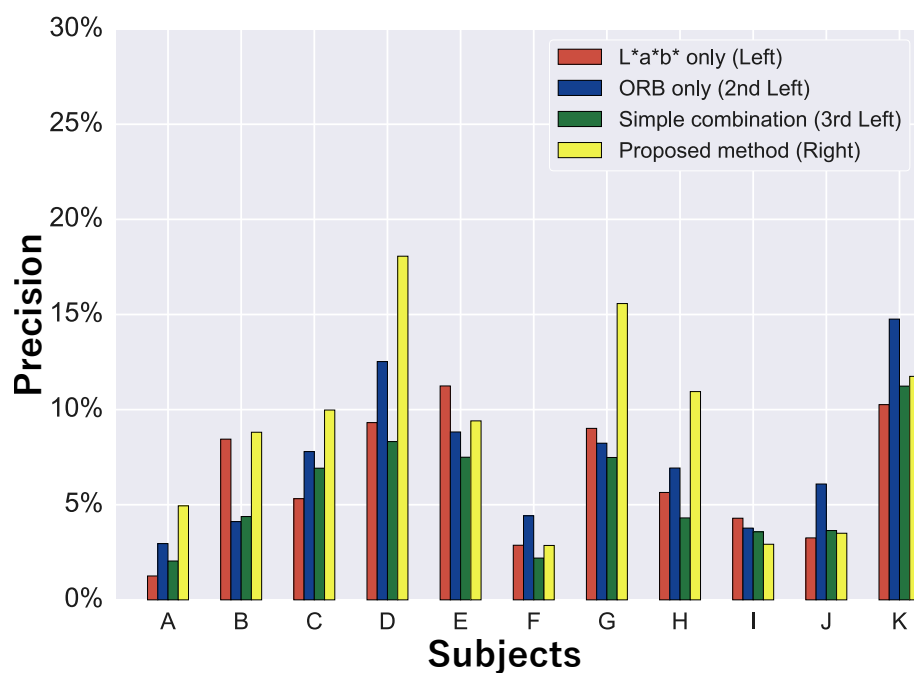


図 4.4 各被験者ごとの適合率

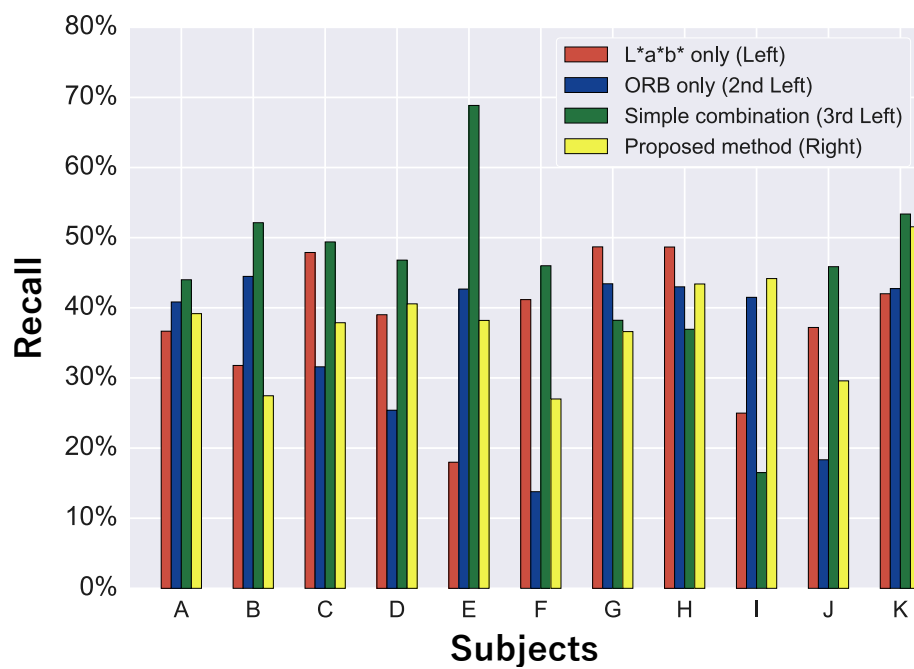


図 4.5 各被験者ごとの再現率

4.3 実験結果

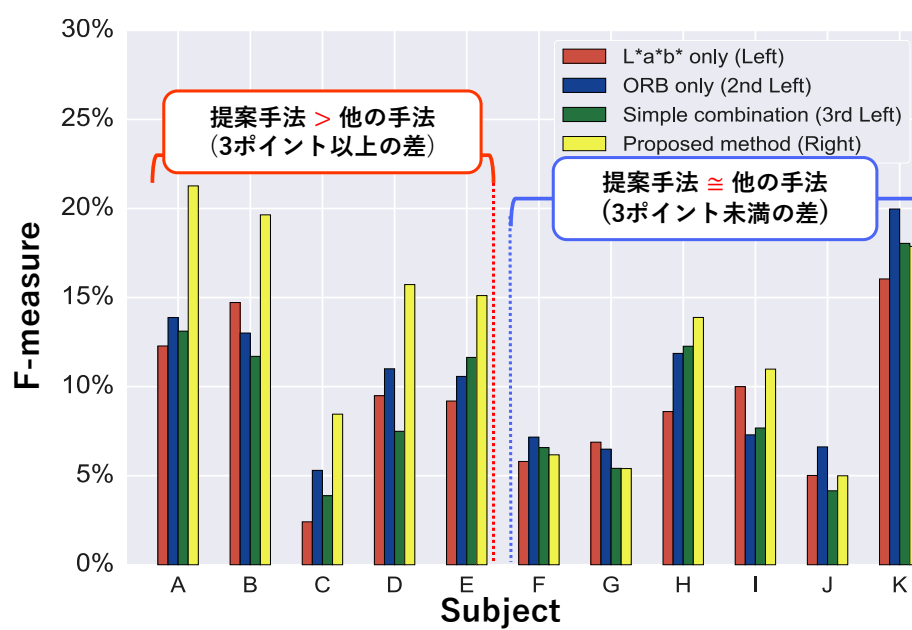


図 4.6 各被験者ごとの F 値

第 5 章

考察

本章では，第 4 章で述べた構築システムの評価に対して考察を行う．

5.1 提示する VK に関する考察

図 3.17，3.18 で示した，提案手法における被験者に提示する VK には類似した画像がある．被験者実験では，どちらの VK が類似しているか悩むことや，両方の VK を選択することが適合率低下の原因の 1 つであると考えられる．



図 5.1 類似する VK

新しいヒストグラムを生成する提案手法 1 では，1 つの VK で色特徴と形状特徴の 2 つを表現するように設計しているため，ユーザは VK の選択が困難なものになる．

5.2 VW 生成に関する考察

本研究では，200 枚の画像，部分画像にして 3200 枚の ORB 特徴は 177805 個抽出され，これをクラス数 50 の K-means 法にかけることで Bag-of-Feature を生成している．この処理時間は約 10 分かかった．これを大規模データセット ImageNet に対応させる場合，

5.3 部分画像クラスタリングに関する考察

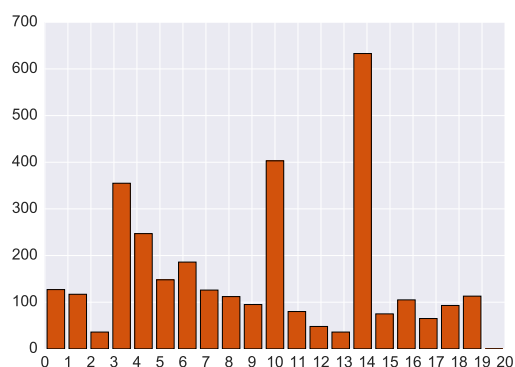
K-means 法の計算量が $\mathcal{O}(n)$ で分散処理が困難なため，ImageNet の全画像 14,197,122 枚の VW 算出にかかる時間は約 492 日かかることになる．そのため，大規模データセットによる Bag-of-Feature の生成には，次元削減や並列処理を用いて，より高速なクラスタリングが要求されると考える．

5.3 部分画像クラスタリングに関する考察

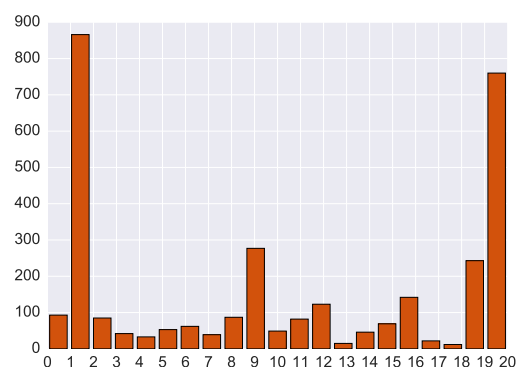
図 5.2 にクラスタリングした全部分画像の分布を示す． $L^*a^*b^*$ 色ヒストグラムと ORB ヒストグラムは最大クラスと最小クラスの要素数には，500 以上の差が生じるほどクラスに偏りが生じている．この原因として，クラスタリング手法が誤っている，データセットの画像に偏りがあるなどが考えられる．前者は，各特徴表現での VK 分類に 10 個のクラスで行っているため，クラスの個数が少ないためクラス分布に偏りが生じたと考える．また，図 5.3 の画像 2 枚は被験者実験において類似していると評価される割合が高かった画像であるが，Art Explosion のカテゴリとしては，異なる画像と定義されている．つまり，データセット内の異なるカテゴリで類似した画像が存在するため，クラスに偏りが生じたと考える．

部分画像のクラスタリングにも K-means 法を用いているが，K-means 法は初期のクラス中心の座標がランダムに配置されるため，毎回クラスタリング結果が多少異なる．Ward 法は初期値依存がないため，毎回正確なクラスタリングを行うことができるが，計算時間が $\mathcal{O}(n^2)$ であるため，データセットの規模が巨大になった場合は，Ward 法が実用的でないのが問題点である．

5.4 データセットに関する考察



[1]L*a*b* 64 次元



[2]ORB 50 次元

図 5.2 VK のクラスに偏りのある特徴表現



図 5.3 Art Explosion で異なるカテゴリに属する 2 枚

5.4 データセットに関する考察

今回、データセットと検索対象画像は、先行研究と同様のものを用いたが、本研究の被験者実験において F 値=0 の回数を各画像ごとにグラフ化したものが図 5.4 である。画像 No.9 が F 値=0 の回数が 0 であるのに対して、画像 No.4, 6 は F 値=0 の回数が 10 以上あるため、データセット内に検索対象画像と類似した画像がない可能性が考えられる。

5.5 画像のサイズ変更に関する課題

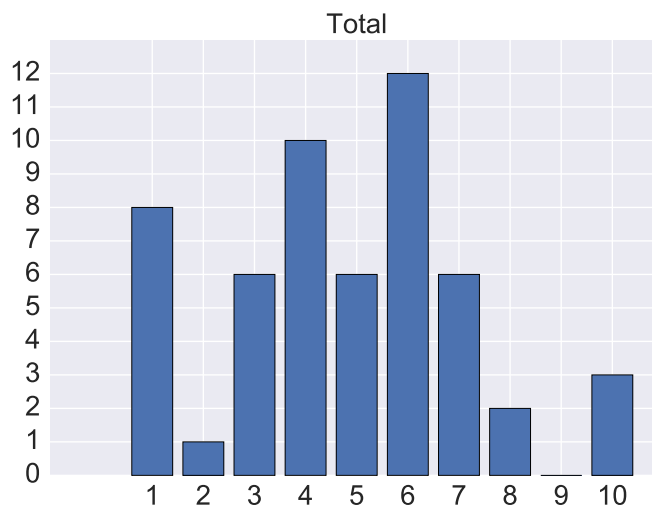


図 5.4 各画像における F 値=0 が算出された回数

5.5 画像のサイズ変更に関する課題

本研究では、部分画像を生成するために元画像を正方形にサイズを変更した。しかし、この変更によって形状が変化してしまいユーザの VK の認識が元画像の形状と異なる可能性がある。そのために、部分画像生成の際に図 5.5 のように元画像のサイズを変更せずに部分画像を生成することができるため、部分画像とユーザ認識の差が縮まると考えられる。

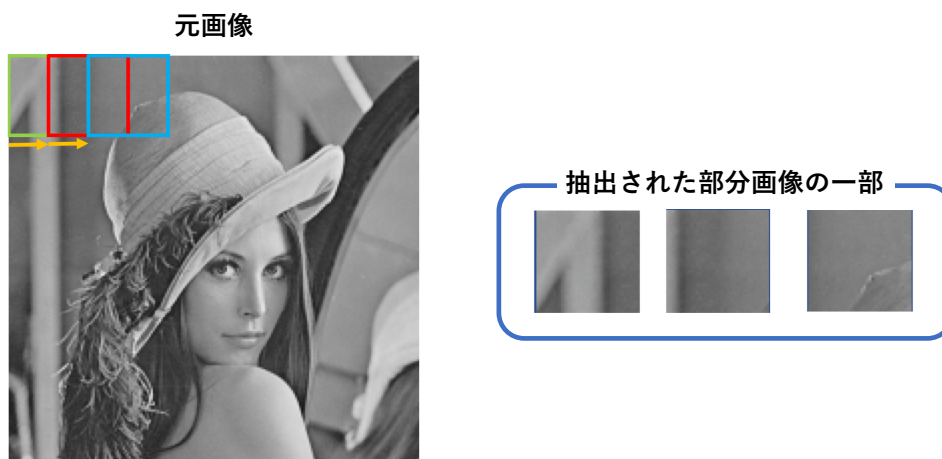


図 5.5 サイズ変更のない部分画像生成

5.6 提示した VK に関する課題

提案手法では，局所特徴と色特徴の各 10 個の VK を提示した．各特徴表現において，選択肢が 10 個しかないのは，クラスタに偏りが生じたり，ユーザが要求する VK が存在しない可能性が高くなる．そのため，各特徴 20 個ずつの VK を提示することで，選択肢が増え，クラスタの偏りが軽減されることで，適合率の向上が期待できる．

5.7 局所特徴のない画像に関する課題

提案手法では，局所特徴の抽出できなかった部分画像はクラスタリングの段階で除いており，検索でも，色の VK と局所特徴の VK の両方を選択することが必須条件となっている．しかし，図 5.6 のような空やシルエットの画像になると，ほとんどの部分画像で局所特徴が検出されない画像が存在する．このような画像は，検索することが困難になる．これに対して，クラスタリングで選定した VK とは別に，局所特徴のない無地の VK を追加することで，局所特徴の持たない画像も検索することができる．



図 5.6 局所特徴が少ない画像の例

5.8 次元削減に関する課題

本研究では、すべての手法において抽出してヒストグラム化した特徴を用いてクラスタリングを行っている。しかし、VW ヒストグラムは、各部分画像から検出された局所特徴の分布なので、不必要に次元が大きい可能性がある。効率的なクラスタリングのため、主成分分析や自己符号化器などを用いて次元削減を行うことで各特徴間の相関がなくなるため、よりクラスタ間に違いが生じることが期待できる。

5.9 データセットに関する課題

本研究で用いた画像データセットは 200 枚の非常に小さなデータセットを用いたため、被験者実験によって、少し VK の選択が異なるだけで、評価値が大きく変動した。画像検索システムの実用性を向上させるために、データセットの大規模化は必須であると考える。しかし、ImageNet の各カテゴリから 1 枚ずつ選択したデータセットを用いた部分画像検索システムは、F 値が 0.01 周辺の値であるため有意差が出にくいと考える。そのため、ImageNet の複数カテゴリから抽出した数千枚の画像データセットを作成することで、より検索システムの評価が有意なものになると考える。

5.10 被験者実験に関する課題

被験者実験では、各被験者は検索対象画像 10 枚を 4 手法で検索、検索結果での適合画像の選択、検索で取れなかった適合画像の選択を各試行で行うため、被験者 1 人あたり約 1 時間の実験時間を要する。そのため、実験の後半で適合画像の選択数の減少などの集中力の低下が感じられた。より被験者の実験時間を短縮するために、ImageNet のカテゴリ定義を用いて適合率、再現率、F 値の算出を検索結果から行うことで、被験者が適合画像を選択する必要がなくなるため、大幅な時間短縮が期待できる。

第 6 章

まとめ

本研究は，色特徴と ORB 局所特徴を用いた，部分類似画像検索における特徴表現方法を提案している．提案する特徴表現方法は，色特徴に従来手法と同様の $L^*a^*b^*$ 色ヒストグラムを用いており，組み合わせ手法として，各部分画像に色特徴と局所特徴をクラスタリングし，2 つの特徴クラスラベルを付与することで各特徴が他の特徴に影響されことなく表現することを可能にしている．

また，各特徴を基に検索を行う部分類似画像検索システムを構築し，被験者 11 人に対して， $L^*a^*b^*$ ，ORB，両ヒストグラムを結合した特徴，提案手法の 4 種類のシステムで検索実験を 10 回ずつ行い，評価として適合率，再現率，F 値を求めた．被験者実験の結果，それぞれの結果 $L^*a^*b^*$ は 0.091，ORB は 0.103，ヒストグラム結合は 0.092，提案手法は 0.127 であった．各被験者の平均 F 値において，提案手法が他の手法より 3 pt 以上高い被験者が 5 名，提案手法と他の手法の差が 3pt 未満の被験者が 6 名，他の手法が提案手法より 3pt 以上高い被験者は 0 名であった．このことから，提案手法が他の手法より有用であると考えられる．

今後の課題としては，画像の分割手法，クラスタリング手法，VK の個数など，本研究では検証しきれなかった要素を検証することで検索精度の向上が期待できる．また，一般的な画像検索に環境を近づけるためにデータセットを再検討する必要があると考える．

謝辞

本研究を進めるにあたり、高知工科大学情報学群の吉田真一准教授には研究活動および様々な面に関して、丁寧に指導していただき心から感謝いたします。吉田先生には、卒業研究において、手法の解説や研究の方針、さらに論文の細かな体裁などを非常に丁寧に指導して頂きました。また、国際学会発表の際にも論文や発表の面倒を見て頂き、学部生の内に、非常に貴重な経験をさせて頂きました。研究以外では、サーバ・ネットワークの運用に関してご指導して頂き、ちょっとした雑談も非常に密度の濃い研究室生活の要因となりました。修士課程の2年間もご指導のほどよろしくお願い致します。

また、高知工科大学情報学群 福本昌弘教授と妻鳥貴彦准教授には、夏合宿などを通じて、複数回ご指導をして頂きました。また、卒業論文の梗概のご指導、研究発表会においても質問をして頂き、研究をより良いものにすることができました。お二人に深く感謝致します。

同期である、佐々木氏、首藤氏、鈴木氏、中山氏、領内氏とは、研究生活の苦楽を共にし、相談や愚痴などを話し合うことで、楽しい研究室生活を送れました。佐々木氏は、同期で唯一の男同士ということで、趣味の話などで研究の息抜きが有意義なものになりました。学業では、いつも土壇場でタスクをこなす集中力、また、研究室内での同期や後輩へのフレンドリーな接し方は見習うものがありました。首藤氏は、独特なキャラクターで周りの雰囲気をおもしろくしてくれるため、研究室の雰囲気は楽しいものでした。また、興味のあるものに対する集中力が高く、研究とサークル活動を両立しているパワフルさは、素晴らしいものだと思います。鈴木氏は、研究テーマが同じで、お互いの研究について相談することもあり、切磋琢磨できたと感謝しております。真面目な性格は、感化されるものがあり、天然な性格は、周りを笑わせてくれたため、とても良い人柄だと思います。中山氏は、卒業研究と大学院受験とアルバイトの両立をするために、研究室で夜遅くまで作業している姿が印象深く、その姿を見て、自分ももう少し頑張ろうという気持ちにさせてくれました。領内氏は、英語が堪能で、留学生とのコミュニケーションや英語の文章作成に関して大変お世話になり

謝辞

ました。また、色々なところに気が回るため、イベント等、様々な場面での的確なアドバイスが非常にありがたかったです。

また、同研究室の3年生皆様には、イベントの幹事やオープンキャンパスなどで活躍して頂きました。3年生が配属されて、研究室の雰囲気が一気に賑やかなものとなり、それぞれ研究室で熱心に作業を行っている姿に刺激を受けました。来年度は、先輩として、卒業研究や就職活動を頑張ってください。

最後に、大学生活を金銭面および精神面で支えて頂いた家族に深く感謝致します。また、大学生活で出会った全ての皆様に感謝致します。

参考文献

- [1] M.Serata, et al. “Designing Image Retrieval System with the Concept of Visual Key,” Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.10, pp.136–144, 2006.
- [2] A.W.M.Smeulders, et al. “Content-based image retrieval at the end of the early years,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.12, pp.1349–1380, 2000.
- [3] 鈴江直人, “モノクロ画像検索のための形状特徴,” 平成 21 年度高知工科大学学士学位論文, 2010.
- [4] 上村香樹, “SURF 特徴を用いた部分画像に基づく類似画像検索,” 平成 25 年度高知工科大学学士学位論文, 2014.
- [5] 竹中浩貴, “ORB 特徴を用いた画像の一部の類似性に基づく画像検索,” 平成 26 年度高知工科大学学士学位論文, 2015.
- [6] 西本高志, “ $L^a \times b$ 色ヒストグラムによる部分領域類似画像検索の ImageNet への応用,” 平成 27 年度高知工科大学学士学位論文, 2016.
- [7] T.Sasatani, et al., “Region-Based Image Retrieval using Local Image Features,” ISCHIA2016, SA-OS-04, Beijing, China, Nov.3-6, 2016.
- [8] E.Rublee, et al. “ORB: an efficient alternative to SIFT or SURF,” Proceedings of the 2011 International Conference on Computer Vision, IEEE Computer Society, 2011.
- [9] 藤吉弘亘, 安倍満, “局所勾配特徴抽出技術 -SIFT 以降のアプローチ-, ” 精密工学会誌, Vol.77, No.22, 2011.
- [10] Richard Szeliski, et al. “コンピュータビジョン アルゴリズムと応用 第 2 版,” 共立出版, pp.70–73, 2014.

参考文献

- [11] “OpenCV 3.2.0-dev Color conversions,”
http://docs.opencv.org/trunk/de/d25/imgproc_color_conversions.html,
2017/1/26 参照.