



記事

帰属関係SIFTに基づく領域グラフ。概念と応用

マリオ・マンゾーニ。

ナポリ "オリエンタル" 大学情報技術サービス、80121 ナポリ、イタリア ; mmanzo@unior.it ; Tel: +39-081-6909229

† 現住所 Via Nuova Marina, 59, 80133 Naples, Italy.

Received (受理). 2020年6月12日; 受理された。2020年8月3日、掲載 :
2020年8月6日

概要：実世界において、構造化されたデータをグラフで表現することが多くなっている。一般に、アプリケーションは最も多様な分野に関係し、データは局所のおよび空間的なつながりの観点から表現される必要がある。このシナリオでは、以前に紹介した **Attributed Relational SIFT-based Regions Graph (ARSRG)** と呼ばれる、デジタル画像の表現のための構造を提供することが目標である。ARSRGはこれまで詳しく説明されておらず、この目的のためには、未知の側面を探ることが重要である。そこで、第一に、上記で規定されていない正式な定義を示し、その構造構成を明らかにする基礎理論を提供すること、第二に、異なるアプリケーションへの適応性と柔軟性に関する重要な要素を提供する実験的な目標を掲げている。理論と実験の組み合わせにより、ARSRGが様々なコンテンツを含む画像の表現に適応可能であることを明らかにする。

キーワード：グラフを用いた画像表現と解析、画像分類、カーネル法、グラフマッチング、グラフ埋め込み、グラフワードのバッグ

1. はじめに

人間の視覚に関する問題の中で、視覚的に複雑な実体の処理は、最も重要なものの一つである。情報処理は、しばしば、ローカルからグローバル、あるいはグローバルからローカルの接続に基づいて行われる[1]。ローカルからグローバルへの概念は、シーンのローカルな詳細からグローバルな構成への遷移に関係し、グローバルからローカルは逆の順序で、グローバルな構成から詳細に向かって動作する。例えば、顔認識のアルゴリズムは、**local-to-global**のアプローチを用いており、目、鼻、耳の認識から始まり、最終的に顔の構成に至ります。一方、グローバル・トゥ・ローカルのアルゴリズムでは、まず顔を認識し、次に細部（目、鼻、耳）を認識することになる。人間の認識作業では、特に被験者が短時間に画像を見ている場合、シーンのグローバルな構成が重要な役割を果たす。さらに、人間はシーンのカテゴリーを認識する効果的な方法として、局所的な情報を活用する。高次の視覚認識理論では、個々の要素をローカルレベルとグローバルレベルに区別し、多くのローカル成分の情報が知覚的にグループ化されている[2]。グラフ表現におけるノードとエッジは、生データ間の関係を強調する目的で情報を符号化する。コンピュータビジョンやパターン認識などの多くの分野で、データグラフ表現とそれに関連する操作アルゴリズムが採用されている。特に、画像処理の分野では、デジタル画像を表現するために多くの方法でグラフが使用されている。標準的な手法では、画像を支配的な不連続領域に分割し、局所の特徴と空間的特徴をそれぞれノードとエッジとして表現する。局所の特徴は、領域の固有特性（形状、色、テクスチャなど）を記述し、空間的特徴は、近傍の位相的情報を提供する。画像表現は、画像検索の分野で働くシステムにとって重要なステップの一つである。現代のコンテンツベース画像検索（CBIR）システムでは、画像全体から抽出された画像の基本要素（色、テクスチャ、形状、位相的關係）を基本的に考慮する。

を効果的に表現するために必要です。これらの要素の分析を通じて、構造的な構造が作り出される。また、領域ベース画像検索[3] (RBIR) と呼ばれるシステムでは、コンテンツ全体ではなく、特定の画像領域に注目し、特徴を抽出する。本論文では、Attributed Relational SIFT-based Regions Graph (ARSRG) と呼ばれる画像表現のためのグラフ構造を説明し、分析し、先行研究 [4-7] を参照しながら議論する。その構成要素の定義による構造の検討と、これまでに得られた結果の収集と分析の2つの主要な部分がある。ARSRGの主目的は、階層的な記述によって、画像の局所的な特徴と大域的な特徴の間のつながりを作ることである。大域的特徴はセグメンテーション技術を用いて抽出され、局所の特徴はLIFE (Local Invariant Feature Extraction) 法に基づいて抽出される。この構造により、画像領域、領域間の位相関係、局所不変特徴から生じる様々な情報を得ることができる。安定した記述子 (変形に頑健な記述子) を抽出するために、LIFE手法の中から、画像のスケール空間における顕著な点を抽出するSIFT特徴を選択した。さらに、構造の詳細な分析から生じる新しい定義と特性を紹介する。この理論的分析は、異なる定義の導入に基づき、ARSRGの主成分と副成分に踏み込み、構築と比較の段階をより良く理解することを目的としている。最後に、幅広い実験段階を通じて、この構造がどのように異なるタイプのアプリケーションコンテキストに適応できるかが示された。後者では、画像コンテンツとアプリケーションの両方の観点から、異なる分野でこれまでに得られた結果の収集と詳細な分析が行われました。これは、主に理論的根拠を支える共通の側面を特定することを目的とした、重要な段階であった。本論文は以下のように構成されている。セクション 2 では、SIFT (Scale-Invariant Feature Transform) [8]を含む、グラフベースの画像表現に関する関連研究を紹介する。セクション 3-5 では、ARSRG の説明、定義、および特性について述べる。セクション6と7では、それぞれ実験結果と結論が報告されている。

2. 関連作品

SIFT特徴量から生じる局所的な情報と空間的な情報を組み合わせるアプローチは、文献に多数報告されている。一般に、グラフ構造は、画像のある位置にあるキーポイントに関する情報を符号化します。ノードはSIFT記述子を表し、エッジは異なるキーポイント間の空間的な関係を記述する。

9]では、グラフ G_1 は、画像 I_1 からSIFTキーポイントの集合を表し、次のように定義される。

$$g_1 = (v_1, m_1, y_1) \quad (1)$$

ここで、 $v_\alpha \in V_1$ は位置 $(p^{(\alpha)}, p^{(\alpha)})$ を持つSIFTキーポイントに関連する 2^{nd} ノード、 $y_\alpha \in Y_1$ はSIFTノード v_α に付属する記述子、 M_1 は隣接行列である。 $M_{1\alpha\beta} = 1$ とすると、ノード v_α と v_β は隣接している場合、 $M_{1\alpha\beta} = 0$ それ以外の場合。

10]では、SIFT 特徴の局所的な情報とグローバルな幾何学的情報とを組み合わせ、特徴量の一致をロバストに推定している。この情報は、グラフ構造を用いて符号化される。

$$g_0 = (v_0, b, y) \quad (2)$$

ここで、 $v \in V_0$ は、SIFT キーポイントに関連付けられたノード、 B は隣接行列、 $B_{v,v}$ = ノード v と v が接続されている場合は1、それ以外は $B_{v,v} = 0$ 、一方 $y_v \in Y$ はノード v に関連付けられたSIFTディスクリプタである。

11]では、ノードが画像グリッドに関連する N 個の画像領域に対応し、エッジが各ノードとその隣接する4個のノードとを接続していた。基本要素は画素ではなく、 x (水平) 方向と y (垂直) 方向に拡張された領域である。ノードはグリッド上の座標で識別される。ノードに関連する空間情報は、インデックス $d_n = (x_n, y_n)$ である。さらに、特徴ベクトル F_n が、対応する画像領域と、次にノードとに関連付けられる。画像は 32×32 画素の重複した領域に分割される。各領域に対して128次元のSIFT記述子が4つ抽出され、連結される。

[12]では、グラフに基づく画像表現として、SIFT特徴、MSER [13]、Harris-affine [14]を挙げている。画像 I^P と I^Q を表す 2 つのグラフ $G^P = (V^P, E^P, A^P)$ と $G^Q = (V^Q, E^Q, A^Q)$ があるとき、 V は抽出された画像特徴であるノードの集合、 E は特徴の空間関係である辺の集合、 A は抽出した特徴に関連する情報である属性の集合である。

[15]では、SIFT特徴をハイパーグラフの形で結合している。ハイパーグラフ $G = (V, E, A)$ は、ノード $v \in V$ 、ハイパーエッジ $e \in E$ 、およびハイパーエッジに関連する属性 $a \in A$ で構成される。ハイパーエッジ e は V からサイズ $\delta(e)$ のノードの部分集合を囲むが、 $\delta(e)$ はハイパーエッジの順序を表している。

[16]では、3次元物体認識へのアプローチが示された。SIFT特徴を利用できるようにし、ロバスト性を向上させるために、グラフマッチングの枠組みが用いられている。標準的な手法とは異なり、テスト画像は離散化や量子化の操作によって有限のグラフに変換されることはない。そして、検出時にテスト画像において連続グラフ空間を探索する。このため、局所カーネルを適用し、画像特徴のインデックス付けを行い、高速な検出を可能にする。

[17]では、シーン認識と位相的SLAMを応用した特徴量のマッチング問題へのアプローチが提案された。この目的のために、シーン画像は特定のデータ構造を用いて符号化される。画像表現は、JSEG [18] アルゴリズムによる画像セグメンテーションと、MSER と SIFT 記述子を組み合わせた方法による不変特徴抽出の 2 段階で構築される。

[19]では、視覚的顕著性に基づく SIFT 特徴を抽出し、物体モデルを構築するために選択された。また、オブジェクトをコンパクトにモデル化するための Class Specific Hypergraph (CSHG)が紹介された。ハイパーグラフは、異なるドロネーグラフで構築される。それぞれは、オブジェクトの単一のプロトタイプ画像を用いて、選択されたSIFT特徴のセットから作成される。このアプローチを用いることで、最小限のオブジェクトビューによってオブジェクトモデルを表現することができる。

[20]の著者らは、SIFTに基づくグラフ構造表現によって、画像を表現する物体認識問題に対する解決策を提供した。SIFTのキーポイントを結ぶ線を用いて、グラフ構造を作成する。 $G = (V, E, X)$ はグラフを表し、集合 E はエッジ、集合 V は頂点、集合 X は関連するSIFTディスクリプタを表す。ノードは、以下の方法で検出されたキーポイントを表す。

は、SIFT アルゴリズムであり、関連するラベルは、128 次元の SIFT 記述子である。エッジ $e_{ab} \in E$ は、2つのノード $u_a \in V$ と $u_b \in V$ を結ぶ。グラフは、画像から抽出されたすべてのキーポイントがそれらの間で接続されている場合、完全であると定義することができる。形式的には、エッジの集合は以下のように定義される。

$$E = \{e_{ij} \mid \forall i, j, \frac{p_i - p_j}{\sqrt{\sigma}} < \lambda\} \quad (3)$$

ここで、 $p = (p_x, p_y)$ はキーポイントの空間座標、 σ はスケール、 λ は閾値である。このようにすることで、余分なエッジを作らないようにする。このように近接グラフを定式化することで、計算量を減らすと同時に、検出性能を向上させることができる。

[21]では、中央値K-最近傍 (K-NN) グラフ $G_P = (V_P, E_P)$ が定義されている。の頂点 v_i は $i \in 1 \sim N$ さらに、 p_j が p_i の K 個の最近傍の一つであり、 $p_i - p_j \leq \eta$ のとき、無向きの辺 (i, j) が生成される。 η は頂点のペア間のすべての距離の中央値で、次のように定義できる。

$$\eta = \text{中央値}_{(l,m) \in V_P \times V_P} \|p_l - p_m\| \quad (4)$$

K-NN グラフ構築の際、構造を支持する K 個の頂点がない場合、頂点 p_i は完全に切断されたものとみなすことができる。グラフ G_P は $N \times N$ の隣接行列 A_P を持ち、 $(i, j) \in E_P$ のとき $A_P(i, j) = 1$ 、それ以外の場合 $A_P(i, j) = 0$ である。

3. 帰属関係SIFTに基づく領域グラフ

本節では、特徴抽出とグラフ構築の2つの主要なステップに基づき、帰属関係SIFTベース領域グラフ (Attributed Relational SIFT-based Regions Graph: ARSRG) を定義する。特徴抽出は、セグメンテーションアルゴリズムによる画像からの関心領域 (ROI) 抽出からなる。次に、画像領域間の空間的関係を記述するための領域隣接グラフ (RAG) [22]を構築する目的で、画像内の接続成分が検出される。同時に、回転、拡大縮小、平行移動、照度変化、射影変換に対する不変性を保証するSIFT [8] 記述子が、原画像から抽出される。グラフ構築は、グラフ構造の構築からなる。ARSRGでは、ルートノード、RAGノード、リーフノードの3つのレベルを区別することができる。第1レベルでは、ルートノードが画像を表し、第2レベルのすべてのRAGノードに接続される。異なる画像領域間の隣接関係は、RAGノードを通して符号化される。したがって、画像中の隣接する領域は、接続されたノードによって表現される。また、各RAGノードは、上位レベルのルートノードに接続されている。最後に、画像から抽出されたSIFTディスクリプタは、リーフノードを通して表現される。第3レベルでは、リージョンベースとリージョングラフベースの2種類の構成が登場する。領域ベースの構成では、キーポイントはその空間座標に基づいて領域と関連付けられる。一方、領域グラフベースの構成では、同じ領域に属するキーポイントがエッジ (空間的隣接性をエンコードする) によって接続されるように記述される。以下、特徴抽出とグラフ構築のステップを詳細に説明する。

3.1. 特徴抽出

3.1.1. 関心領域抽出

JSEG [18]と呼ばれるセグメンテーションアルゴリズムによって、画像からROIを抽出する。JSEG は、色の量子化と空間的なセグメンテーションという 2 つの異なるステップを経てセグメンテーションを行う。最初のステップは、画質を大きく劣化させることなく、粗い量子化を行う。第二段階は、関連する画素の色の類似性を考慮せずに、クラスマップ上の空間的なセグメンテーションを行う。

3.1.2. 接続部品へのラベル付け

次のステップでは、セグメンテーションの結果に対して、連結成分のラベリングを行う。連結成分とは、同じ色の連続した画素で構成される画像領域のことである。画像Bに連結成分のラベリングを行う処理により、ラベル (正の整数または文字) を含む出力画像LBが生成される。ラベルとは、ある実体を排他的に名付ける記号のことである。4近傍と8近傍で接続された領域は同じラベルを持つことになる (アルゴリズム1、2では、数値を含む変数 m で表現されている)。アルゴリズム1は連結成分のラベリング版を示している。

アルゴリズム1 連結成分ラベリング。

要求する。I - イメージからラベルへ。

Ensure: I - Image Labeled;

1: $m=0$

2: **for** $y=1:I_size_y$ **do**

3: **for** $x=1:I_size_x$ **do**

4: **if** $I[i][j] == 0$ **then**

5: $m=m+1$

6: コンポーネントラベル(I, x, y, m)

7: **end if**

8: **end for**

9: **end for**

10: Iを返す

アルゴリズム2 コンポーネント・ラベル

要求する。I - 画像からラベルへ; i, j - 画像インデックス; l - ラベル;

エンシュア : 0;

```

1: if I[i][j] == 0 then
2:   I[i][j]=mとする。
3:   コンポーネントラベル (I, -i1, j1, m)
4:   コンポーネントラベル (I, i1, j, m)
5:   コンポーネントラベル(I, i1, j+1, m)
6:   コンポーネントラベル(I, i, j1, m)
7:   コンポーネントラベル(I, i, j+1, m)
8:   コンポーネントラベル(I, i+1, j1, m)
9:   Component Label(I, i+1, j, m)
10:  Component Label(I, i+1, j+1, m) (コンポーネントラベル)
11: end if

```

3.1.3. 領域隣接グラフ構造

ARSRGの第2レベルは、RAG (Region Adjacency Graph) と呼ばれるグラフベースの画像表現をホストしている[22]。アルゴリズム3では、RAGアルゴリズムの疑似コード版を示している。領域とは、画像セグメンテーションの結果に基づき、画像の素片を表す。RAGでは、ノードは領域であり、エッジは2つのノード間の隣接を記述する。RAGは、領域間の空間的な関係を参照して構築される。2つの領域は、同じ境界を共有する場合、空間的に近いと定義される。このため、2-4行目で $pixel(i, j)$ と名付けたラベル付き領域の画素のラベルと、その近傍（8近傍に含まれる8方向）に属する画素のラベルの間で、近傍チェックを行う。後者が $pixel(i, j)$ に対して異なるラベルを含む場合、RAGに表現される2つの隣接領域が存在することを意味する。RAGはグラフ $G = (V, E)$ として定義され、ノードはV内の領域、エッジEはそれらをつなぐ境界線である。Gは、グラフの接続のトポロジーを記述する隣接行列Adjacency_matrix（5行目）を通して符号化される。例えば、Adjacency_matrix(i, j)が1を含む場合、これは画像上で領域i, jが接続されることを意味する。さらに、RAGの主な特性として、並進や回転に対する不変性があり、高度な画像表現に有用である。

アルゴリズム3 領域隣接グラフ。

要求する。 Labeled_image。

Ensure: グラフ構造(Adjacency_matrix)。

```

1: Adjacency_matrix = 0
2: for pixel(i, j) ∈ Labeled_image do
3:   for pixel(x, y) ∈ 8-neighborhood do.
4:     if pixel(i, j) ≠ pixel(x, y) then
5:       Adjacency_matrix(pixel(i, j), pixel(x, y)) = 1
6:     end if
7:   end for
8: end for
9: return Adjacency_matrix

```

3.1.4. スケール不変の特徴変換

SIFT [8] 記述子は、画像記述における回転、拡大縮小、並進、部分的な照度変化、射影変換に対する不変性を確保するために抽出される。SIFTは、特徴抽出の段階で、RAGの作成と並行して計算される。

3.2. グラフ構造

ARSRGの構築プロセスは、3つのレベルの作成で構成されています。

1. **ルートノード**：グラフ構造の第一階層に位置するノードで、画像を表す。次のレベルのすべてのノードに接続されている。
2. **Region Adjacency Graph (RAG) ノード**：セグメンテーション結果に基づく、異なる画像領域間の隣接関係。したがって、隣接する画像領域は、このレベルで接続されたノードで表現される。
3. **リーフノード**。画像から抽出された SIFT 特徴の集合。2種類の接続が用意されている。
 - (a) **領域ベース**。リーフノードは、特徴抽出の際に得られたSIFTキーポイントを表す。各リーフノードキーポイントは、画像内の空間座標に基づく領域と関連付けられている。このレベルでは、各ノードは1つのRAG上位ノードにのみ接続される（図1a）。
 - (b) **領域グラフベース**。従来の構成に加え、同一領域に属するリーフノード-キーポイントを、閾値基準に基づき、空間的隣接性をコード化したエッジで接続する（図1b）。

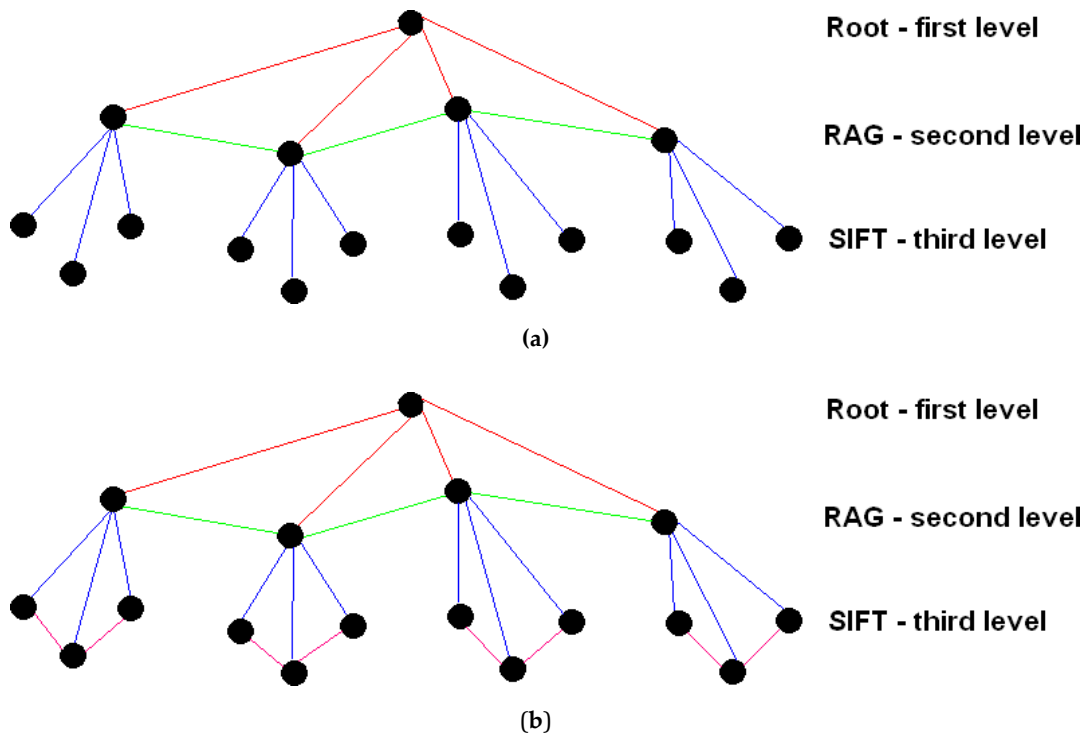


図1.リージョンベース(a)とリージョングラフベース(b)の構成。RAG, Region Adjacency Graph.

4. 正式な定義

ここでは、ARSRGの構造を正式に固定化する目的で、詳細な定義を紹介する。定義1と2は、（図1に示すような）2つの構成に関連するコンポーネントを記述する。定義3、4は、定義5、6で導入した機能を通じて、第2、第3レベルのノードに関連する属性のセットを定義している。定義7と8は、SIFTのための接続構造を導入している。定義9から11では、レベル間の様々なタイプのエッジが記述されている。最後に、定義12と13は、第2レベルと第3レベルのためのサポート構造を含む。ARSRG構造は、2つのリーフノードの構成に基づいて定義されている。

定義1. ARSRG_{1st} (first leaf nodes'

configuration): G はタプル $G =$
($V_{regions}$, $E_{regions}$, V_{SIFT} ,

$E_{regions-SIFT}$) として定義され、ここで。

- $V_{regions}$, リージョンノードの集合。
- $E_{regions} \subseteq V_{regions} \times V_{regions}$ は無向辺の集合で、 $e \in E_{regions}$ 、 $e = (v_i, v_j)$ はノード v_i , $v_j \in V_{regions}$ を結ぶ辺である。
- V_{SIFT} , SIFT ノードの集合。
- $E_{regions-SIFT} \subseteq V_{regions} \times V_{SIFT}$ 、有向辺の集合で、 $e \in E_{regions-SIFT}$ 、 $e = (v_i, v_{fj})$ は、送信元ノード $v_i \in V_{regions}$ と送信先ノード $v_{fj} \in V_{SIFT}$ を結ぶ辺である。

定義2. ARSRG_{2nd}

(second leaf nodes' configuration): G はタプル $G = (V_{regions}, E_{regions}, V_{SIFT}, E_{regions-SIFT}, E_{SIFT})$ として

定義さ、ここで、 $V_{regions}$, $E_{regions}$, V_{SIFT} , $E_{regions-SIFT}$, E_{SIFT} は、
タプルに

- $V_{regions}$, リージョンノードの集合。
- $E_{regions} \subseteq V_{regions} \times V_{regions}$, $e \in E_{regions}$, $e = (v_i, v_j)$ はノード $v_i, v_j \in V_{regions}$ を結ぶ辺である無向辺の集合
- V_{SIFT} , SIFT ノードの集合。
- $E_{regions-SIFT} \subseteq V_{regions} \times V_{SIFT}$ 、有向辺の集合で、 $e \in E_{regions-SIFT}$ 、 $e = (v_i, v_{fj})$ は、送信元ノード $v_i \in V_{regions}$ と送信先ノード $v_{fj} \in V_{SIFT}$ を結ぶ辺である。
- $E_{SIFT} \subseteq V_{SIFT} \times V_{SIFT}$ 無向辺の集合で、 $e \in E_{SIFT}$ 、 $e = (v_{fi}, v_{fj})$ はノード $v_{fi}, v_{fj} \in V$ をつなぐ辺である。SIFT

ARSRG 構造の第一葉および第二葉のノード構成は、定義 1 および定義 2 に基づいて作成されている。セット $V_{regions}$ と V_{SIFT} に属するノードは、画像から抽出された特徴量と関連付けられている。具体的には

定義 3. $F_{regions}$ は、 $V_{regions}$ のノードに関連するベクトル属性の集合である。要素 $f_i \in v_i$ は、第二レベルの ARSRG 構造のノードに関連づけられる。これは、領域の寸法（ピクセル）を含んでいる。

定義 4. F_{SIFT} は、 V_{SIFT} のノードに関連するベクトル属性の集合である。要素 $f_j \in v_{fj}$ は、第3レベルの ARSRG 構造のノードに関連付けられる。これは、SIFT 記述子を含む。

機能とノードの関連付けは、以下のように定義された代入関数によって行われる。

定義5. ノードラベリング関数 $L_{regions}$ は、第2レベルの ARSRG の各ノード $v \in V_{regions}$ にラベルを割り当てる。ノードラベルは画像から抽出された特徴属性 d_i である。ラベルの値は、領域の次元（ピクセル数）である。 v ノードのラベル付け手順は、ARSRG 構築の過程で発生する。

定義6. SIFT ノードラベリング関数 L_{SIFT} は、第3レベルの ARSRG の各ノード $v_{fj} \in V_{SIFT}$ にラベルを割り当てるものである。ノードラベルは、画像から抽出された特徴ベクトル f_j 、キーポイントである。 v_{fj} のノードのラベル付け手順では、キーポイントが属する領域と比較して、画像内の位置を確認する。

さらに、RAG のノード $e \in V_{regions}$ は、それらの間で横方向に、ノード $e \in V_{SIFT}$ と縦方向に二重にリンクしている。エッジ $e \in E_{regions}$ はすべて左から右への無向性であり、エッジ $e \in E_{regions-SIFT}$ はすべて上から下への有向性である。ルート・ノードは RAG ノードに向かうエッジのリストを保持する。さらに、各 RAG ノードは、RAG ノードから出力されるエッジ、リーフノードから出力されるエッジ、およびルートノードに入力されるエッジの3つのリンクリストを保持する。最後に、各リーフノードは、RAG ノードへのインゴーイングとリーフノードからのアウトゴーイングの2つのエッジのリンクリストを保持しています。各リストのエッジは、端点ノード間の距離に基づいて順序付けされている。

エッジが最初に来ます。これらの辺のリストは、幾何学的に直接的な意味を持ちます。各ノードは、左、右、上、下という1方向で他のノードと接続されています。

ARSRGの第3レベルの組織については、非常に重要な側面があります。そのために、SIFT Nearest-Neighbor Graph (SNNG)を導入しています。

定義7. $SNNG = (VF_{SIFT}, E_{SIFT})$ は次のように定義される。

- VF_{SIFT} : SIFTキーポイントに関連するノードのセット。
- E_{SIFT} : エッジの集合。ここで、各 $v_i \in VF_{SIFT}$ に対して、 $dist(v_i, v_{ip}) < \tau$ の場合にのみエッジ (v_i, v_{ip}) が存在する。 $dist(v_i, v_{ip})$ は画像中のキーポイントの x, y 位置に適用したユークリッド距離、 τ は閾値、 p は1から k までのステム、 k は VF_{SIFT} のサイズである。

この表記法は、マッチング段階において非常に有用である。実際、各SNNGは、定義2を参照して、画像領域に属するSIFT特徴の集合を示し、局所のおよび空間的な観点から整理されたSIFT特徴を表している。SNNGの異なるバージョンは、完全SIFT最近傍グラフ (SNNGc) と呼ばれる。

定義8. $SNNGc = (VF_{SIFT}, E_{SIFT})$ は次のように定義される。

- VF_{SIFT} : SIFTキーポイントに関連するノードのセット。
- E_{SIFT} : エッジの集合で、各 $v_i \in VF_{SIFT}$ に対して、 $dist(v_i, v_{ip}) < \tau$ の場合にのみエッジ (v_i, v_{ip}) が存在します。 $dist(v_i, v_{ip})$ は画像中のキーポイントの x, y 位置に適用したユークリッド距離、 τ は閾値、 p は1から k の範囲で k は VF_{SIFT} の大きさ、となります。この場合、 τ はキーポイント間の最大距離よりも大きくなります。

もう一つの重要な点は、ARSRG構造におけるノード間の垂直方向と水平方向の関係の違いである。以下では、これらの関係であるエッジを定義する。

定義9. 領域水平辺 $e, e \in E_{regions}$ は、ノード $v_i, v_j \in V_{regions}$ を結ぶ無向辺 $e = (v_i, v_j)$ である。

定義10. SIFT 水平エッジ $e, e \in E_{SIFT}$ は、ノード $v_{fi}, v_{fj} \in V_{SIFT}$ を結ぶ無向エッジ $e = (v_{fi}, v_{fj})$ である。

定義11. 垂直エッジ $e, e \in E_{regions-SIFT}$ は、送信元ノード $v_i \in V_{regions}$ および $v_{fj} \in VF_{SIFT}$ から送信先ノード v_{fj} へ接続する有向エッジ $e = (v_i, v_{fj})$ である。

このように、水平エッジと垂直エッジは、それぞれ同じレベル、異なるレベルのノードを結んでいる。最後に、これらの関係は以下に定義する隣接行列によって表現される。

定義12. 二領域隣接行列 $S_{regions}$ は、RAGノード間の空間的關係を記述する。要素 s_{ij} は、ノード $v_i, v_j \in V_{regions}$ を接続するエッジ $e = (v_i, v_j)$ を定義する。したがって、要素 $s_{ij} \in S_{regions}$ は、ノード v_i がノード v_j に接続されている場合は1、そうでない場合は0に設定される。

定義13. バイナリSIFT隣接行列 S_{SIFT} は、リーフノード間の空間的關係を記述する。要素 s_{ij} は、ノード $v_{fi}, v_{fj} \in VF_{SIFT}$ を接続するエッジ $e = (v_{fi}, v_{fj})$ を定義する。したがって、要素 $s_{ij} \in S_{SIFT}$ は、ノード v_{fi} がノード v_{fj} に接続されている場合は1、それ以外は0とする。

図2は、サンプル画像上の2種類のARSRG構造を示したものである。

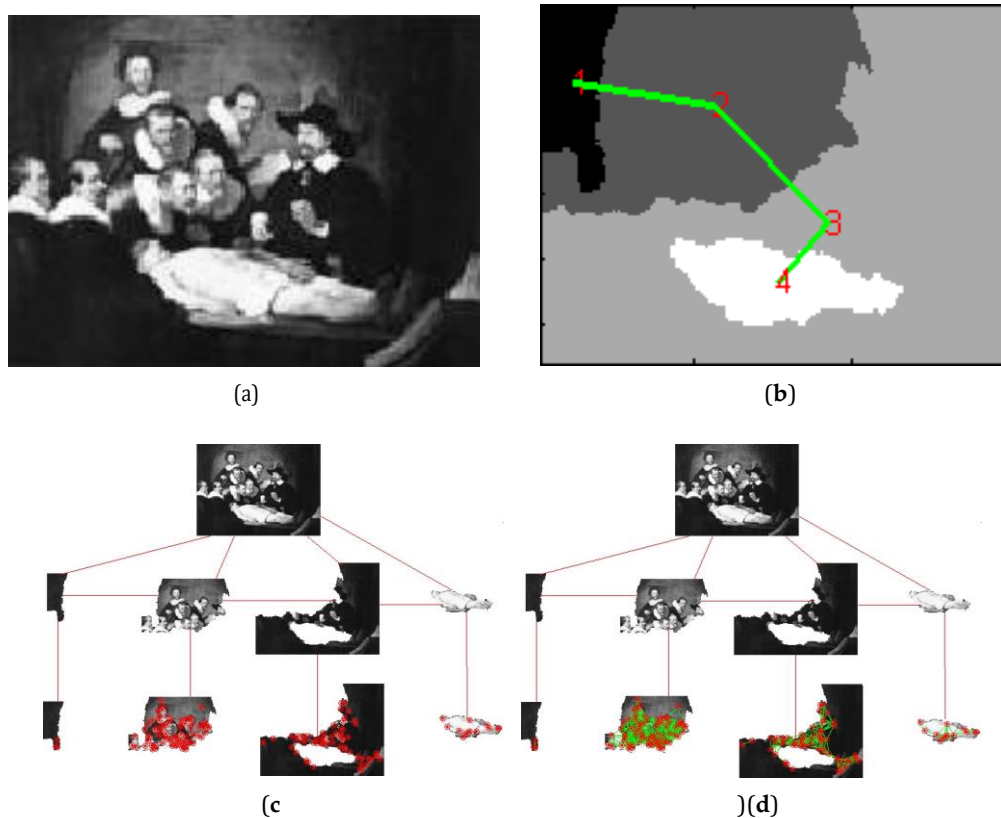


図2 (a) 原画像、(b) 4つの領域からなるRAG、(c) 領域ベースのリーフノード構成、(d) 領域グラフベースのリーフノード構成。(c,d)の赤い点は領域に属するSIFTキーポイント、(d)の緑の線はグラフベースのリーフノード配置のエッジを表す。

5. 物件紹介

ここでは、特徴抽出とグラフ構築のステップから生じるARSRG構造の特性について説明する。

領域の特徴と構造情報ARSRGの構造の主な目的は、次のように接続することです。地域的な特徴や構造的な情報最初のステップは、ROI を抽出するための画像セグメンテーションである。これは、シーンから意味情報を抽出するためのステップである。画像が分割されると、RAG構造が作成される。この特徴量表現は、個々の領域とその間の空間的な関係を強調する。

水平方向と垂直方向の関係ARSRGの構造は、画像特徴間の関係（エッジ）として、水平方向と垂直方向の2種類を提示する。垂直方向のエッジは画像のトポロジー構造を定義し、水平方向のエッジはノード（領域）特徴の空間的制約を定義する。水平方向の関係（定義9と10）は、構造の第二階層に位置するROIとSIFT特徴に関係する。一般的な目的は、空間的な近さの情報を提供し、ノード属性の空間的な制約を定義し、定義された画像上の特定の解像度レベル（詳細）の特徴マップを特徴付け、計算の複雑さと発生頻度に従って区別することができることである。その順序は、 $\{1, \dots, n\}$ の範囲にあり、ここで n は関係である。別の言い方をすれば、垂直関係（定義11）は、個々の領域とその特徴を示す。垂直方向の有向辺は、ARSRGの第2レベルと第3レベルの間のノード（RAGノードからリーフノード）を接続し、親子関係を提供する。この文脈では、ARSRG構造の役割は、定義された関係間のブリッジを作成することである。この側面は、マッチングプロセスにおいて、構造を広く、深く探索することが可能であるという利点をもたらす。

視点、照明、スケールに対して不変な領域特徴量。局所不変な領域記述子の構築は、物体認識、マッチング、再構成などの応用が期待されるホットな研究テーマである。ここ数年、ある種の幾何学的・写真的変換に対して不変な記述子の設計に大きな成功が得られている。局所不変特徴抽出 (LIFE) 法は、画像の特徴的な領域の集合から、安定した記述子を抽出するために利用される。LIFE法は、特定の条件に対する不変性を提供するために、領域表現に選択された。各領域から抽出された情報を用いて作成された局所表現は、照明や視点変更などの特定の画像変形に対してロバストである。ARSRG構造は、異なるLIFE手法の間で最も安定した表現として[23]で同定されたSIFT特徴を含んでいる。

異なるレベルに配置された詳細な情報による優位性。ARSRGの構造によって提供される詳細な画像記述は、比較段階での利点を表している。階層的な方法で、マッチング手順は、ARSRG内のグローバル、ローカル、および構造的な情報を探索する。最初のステップは、サイズに基づく領域のフィルタリング手順である。情報の乏しい小さな領域は除去される。その後、マッチング手順はARSRG構造の次のレベルに進み、より強いマッチングを得るために、単一領域の特徴を分析する。目標は、ARSRGの複数のSNNG (定義7) 上のマッピングを解くことである。本質的に、この基準は、ARSRGsに属するSNNG間の部分的なマッチングを識別する。この手順の間、グラフSNNGの異なる組み合わせが識別され、マッチングプロセスの階層が構築される。このようにして、全体の複雑さが軽減され、特に大きなARSRGに対して大きな利点を示すことが期待される。

領域単位でのマッチングによる優位性領域ベースの画像検索 (RBIR) [24]システムは、領域の特徴に基づいて2つの画像間の類似性を抽出し定義するという目標で動作します。ユーザーは、画像全体ではなく、特定の領域に注意を向けることが実証されている。領域ベースの画像表現は、より人間の知覚に近いことが証明されている。そこで、ARSRGの構造を比較するために、画像分割結果の外観の類似性に基づく領域マッチング方式を採用することができる。領域マッチングアルゴリズムでは、セグメンテーションによって得られた領域を利用し、それらに関連する特徴を比較する。一対の領域の類似度は、領域に属するSIFT特徴量の集合から計算される。マッチング手順は非対称的である。入力画像は領域に分割され、そのSIFTキーポイントのグループは、他の画像の一貫した部分内でマッチングさせることができる。このように、セグメンテーションの結果を利用して、キーポイント候補の領域を作成することで、同じシーンの2つの画像に対して、互換性のない領域を回避することが可能である。

疑似マッチの除去 LIFE法の主な課題の1つは、誤判定の除去である。の誤マッチングが発生します。LIFE法は、比較の段階で、精度に大きな影響を与える多くの偽のマッチングを生成することが示されている。主な理由は、画像特徴間の対応性の欠如 (例えば、シーンの部分的な背景隠蔽のため) である。特徴量の記述子に基づく標準的な類似度測定は、たとえそれが領域の外観にのみ依存しているとしても、広く使用されています。場合によっては、正しいマッチングを確保するために十分な識別ができないことがある。この問題は、テクスチャが少ない場合や均一な場合に関連性が高くなり、多くの誤マッチングを引き起こす。ARSRG構造の適用は、この問題の解決策を提供する。誤マッチングを減らすために、小さなARSRG領域ノードと関連するSIFTディスクリプタは削除されます。実際、小さな領域とそれに関連する特徴は、画像の説明やマッチングにおいて、あまり有益ではありません。比率検定[8]またはグラフマッチング[25]を適用して、残りの領域間の比較を行うことができます。このフィルタリング手順は、実験に強い影響を与え、結果として適切な精度向上をもたらします。

6. 実験結果

このセクションでは、さまざまな応用分野から得られた実験結果を紹介します。特に

1. グラフマッチング[4]。ARSRGは、芸術絵画の検索問題に対処するために採用される。ARSRGの類似性は、局所情報とトポロジー関係を利用し、グラフマッチングアルゴリズムによって計測される。
2. グラフ埋め込み[5]。物体認識問題に効果的に取り組むために、ARSRGを採用する。グラフ構造をベクトル空間に埋め込む枠組みが構築されている。
3. Bag of graph words [6]と呼ばれる。分類タスク内の画像表現に、ARSRGの頻度であるヒストグラムを符号化したベクトルを採用。
4. カーネルグラフ埋め込み[7]。不均衡な分類問題に効果的に取り組むために、ARSRGを採用する。KGEARSRG (Kernel Graph Embedding on Attributed Relational Scale-Invariant Feature Transform-based Regions Graph) は、ベクトルベースの画像表現を提供するものである。

6.1. グラフマッチング

このセクションでは、以前[4]で得られた結果をレビューする。ARSRGをLIFE法、グラフマッチングアルゴリズム、CBIRシステムと比較するために、3つのデータセットが採用されている。最初のデータセットは[26]に記述されており、Olga's gallery (<http://www.abcgallery.com/index.html>) とTravel Webshots (<http://travel.webshots.com>) の画像を統合して得られたものである。27]のデータセットは、Cantor Arts Center (<http://museum.stanford.edu/>)で撮影された絵画の写真から得られたものである。3番目のデータセットは[28]に記述されているもので、1002枚の画像が含まれている。図3はいくつかの例を示している。



図3.(a,b)では、アートペインティングの例をいくつか報告している。

ディスカッション

LIFE手法は、[26]で採用されたデータセットで、Mean Reciprocal Rank (MRR)の観点から比較されている。時間的特徴に基づく実験であるため、この分野の典型であるマッチング段階での偽陽性問題に対する解決策を見つけることが目的である。表1は、ARSRGが到達した最高の結果を示しています。この改善は、特徴間のトポロジカルな関係と、画像から抽出された特徴の完全なセットに対するフィルタリングに関連しています。実際、多くの偽のマッチングを破棄する目的で、標準的なアプローチで提案されているように、画像全体ではなく、領域に属するディスクリプタが比較されます。グラフマッチングと比率検定の両方において、誤判定の許容度を制御する ρ パラメータの最適値は、チューニングにより見出される。

8,26]に記載されているように、(26)では0.6と0.7が受け入れられ、[8]では0.8以上の値が拒否された。ρ値は、ARSRGマッチングでは0.7と0.8が最適です)。

表 1.26]のデータセットにおけるSIFT [8], SURF [29], ORB [30], FREAK [31], BRIEF [32], Attributed Relational SIFT-based Regions Graph (ARSRG) matchingのMRR (Mean Reciprocal Rank) 計測による量的比較。

ρ SIFTSURFORB	フリー ク	BRIEF	アールエ スアール ジーワン スト	ARSRG2n d
0.60.74850.84000.6500	0.3558	0.4300	0.6700	0.6750
0.70.70510.68000.6116	0.3360	0.3995	0.7133	0.7500
0.80.69630.59970.5651	0.2645	0.4227	0.6115	0.8000

27]のデータセットでの性能を測定するために、プレジジョンとリコールが採用されている。表2から分かるように、SIFTに基づくアプローチは、想起の点で優れている。また、ARSRGマッチングでは、ρが0.8であれば、同等の結果が得られる。一方、表3では、ρが0.6、0.7、0.8の場合、ARSRGマッチングが精度において最良のアプローチである。この結果は、画像構造表現を適用した結果として得られたものである。実際、ARSRGのノードである画像領域は、SIFTセット全体にわたる分割ルールを提供する。このようにして、処理中に部分集合が別々に選択される。この戦略は、最終的なランキングの候補としていくつかの画像を破棄する目的で、ほとんどの偽のマッチングを削除する。

表2.27]のデータセットにおけるSIFT [8], SURF [29], ORB [30], FREAK [31], BRIEF [32] およびARSRGマッチングの定量的比較（想起尺度を使用）。

ρ	シフト部	SURFORB	フリー ク	BRIEF	アールエ スアール ジーワン スト	ARSRG2n d
0.6	1.0	0.86660.8000	0.7333	0.7666	0.7333	0.7333
0.7	1.0	0.90000.8666	0.7333	0.8666	0.7666	0.7333
0.8	1.0	1.01.0	0.8333	1.0000	0.8000	0.8000

表3.27]のデータセットにおけるSIFT [8], SURF [29], ORB [30], FREAK [31], BRIEF [32], ARSRGマッチングの精度指標による定量的な比較。

ρ SIFTSURFORB	フリー ク	BRIEF	アールエ スアール ジーワン スト	ARSRG2n d
0.60.06740.08200.2051	0.05584	0.10689	1.0	1.0
0.70.04010.04410.0742	0.04671	0.05664	1.0	1.0
0.80.03120.03380.0348	0.04072	0.03452	1.0	1.0

グラフSIFTに基づくマッチングアルゴリズム（HGM [15], RRWGM [33], TM [34]）の比較には、[26,28]に記述されたデータセットを採用した。MRRの結果を表4と表5に示す。ここでも、ARSRGは領域マッチングのアプローチを採用することで、より良い結果を得ることができた。これは、偽のマッチングを除去し、最終的な結果を向上させる。グラフマッチング問題の重要な点は、比較する画像で表される異なる集合に属する、関連付けるべきノード間の対応規則に関するものである。標

準的なケースでは、2つの集合全体から選択される。ARSRGは、画像全体から抽出された特徴量の主要な集合を分割・間引きすることで、性能と実行時間の改善を実現します。

表4. 26]のデータセットにおけるHGM [15], RRWGM [33], TM [34]アルゴリズムとARSRGマッチングのMRR指標を用いた量的比較。

HGM	RRWGM	TM	ARSRG1st	ARSRG2nd
0.26000	0.13220	0.13480	0.6115	1.0

表5.28]のデータセットにおけるHGM [15], RRWGM [33], TM [34]アルゴリズムとARSRGマッチングのMRR指標を用いた量的比較。

HGM	アールダ ブリュー ジーエム	TM	アールエ スアール ジーワン スト	ARSRG2nd
0.1000	0.0545	0.0545	0.20961	0.39803

表 6 は、Lucene Image Retrieval (LIRe) [35] システムと関連する特徴量との、[26] で示されるデータセット上での性能比較と MRR の観点からの比較を記述したものである。MPEG7 [36], Tamura [37], CEDD [38], FCTH [39], ACC [40]である。LIReの性能の低さから、LIReは絵画検索に適さないことがわかる。この挙動は、関連画像と非関連画像の識別がうまくいかず、最終的なランキングに不適切な結果が含まれていることに起因する。一方、ARSRGはこの問題に対して非常に適している。確かに、ARSRGに含まれる局所的な情報と構造的な情報の2つの情報が、パフォーマンスを向上させる要因になっています。このように、画像の内容は、その部分の関連性に関して記述されます。そうでなければ、局所的な情報しか提供しないLIReが採用する機能で起きてしまいます。

表 6.Lucene Image Retrieval (LIRe) [35] システムで利用可能ないくつかの特徴と、[26] のデータセットにおける ARSRG マッチングとの、MRR メジャーを用いた定量的な比較。

MPEG7	Tamura	CEDD	FCTH	ACC	ARSRG1st	ARSRG2nd
0.7500	0.26450	0.18850	0.23290	0.19240	0.18790	0.7133

6.2. グラフエンベディング

本節では、以前に[5]で得られた結果をレビューする。ARSRG をテストするために採用した、サイズ、デザイン、トピックが異なる 3 つのデータセットについて、以下に説明する。

1. コロンビア画像データベースライブラリ (COIL-100) [41]は、100個のオブジェクトから構成されている。各物体は、それを異なる回転視点で示した72枚のカラー画像で表現されている。物体は黒い背景の上に配置されている。
2. Amsterdam Library Of Images (ALOI) [42]は、1000個の小物体からなるカラー画像集である。COIL-100では、画像全体を埋め尽くすように物体が切り取られているのに対し、ALOIでは、画像に背景と物体がそのままの大きさで含まれている。物体は黒い背景の上に配置されている。
3. ETH-80 [43]は、8つのカテゴリからなる80の物体から構成され、各物体は41の異なるビューで表現されているため、合計3280枚の画像が得られている。物体は一律な背景の上に配置されている。

図4は、オブジェクトの例である。



図 4.Columbia Image Database Library (COIL-100) データセット (a,b), Amsterdam Library Of Images (ALOI) データセット (c,d), ETH-80 データセット (e,f) の画像例です。

ディスカッション

ETH-80データベースと[44]に関連するセットアップで得られた結果を表7にまとめている。学

習セットは240枚の画像からなり、各カテゴリ（リンゴ、車、牛、カップ、馬、トマト）に対して、4つのオブジェクトと、各オブジェクトに対して、10種類のビューから構成される。テストセットは

の残りの画像、1カテゴリあたり60枚（1オブジェクトあたり15ビュー）。結果は、Logistic Label Propagation (LLP) [45] + Bag of Words (BoW) [46]) と [47] (gdFil), [48] (APGM), [49] (VEAM) のアプローチを適用し、[44] で述べたものである。表7では、LLP分類器を採用した ARSRG埋め込みが最も良い性能を示している。この場合、ARSRG埋め込みがオブジェクトビューの変更に正しく対処できることが確認できる。

表7.ETH-80データベースにおける認識精度。

方法	精度
LLP+ARSRGem b	89.26%
LLP+BoW	58.83%
gdFil	47.59%
エーピーエム	84.39%
VEAM	82.68%

[44,50]に関連する設定により COIL-100 データベースで得られた結果を表8にまとめる。特に、25個のオブジェクトのうち11%をランダムに選択したものを学習セットとし、残りのものをテストセットとする。結果は、Logistic Label Propagation (LLP) [45] によって達成されている。
+ また、[44,50]では、彼らの手法（VFSR）と[47]（gdFil）, [48]（APGM）, [49]（VEAM）, [51]（DTROD-AdaBoost）, [52]（RSW+boosting）, [53]（sequential patterns）, [54]（LAF）が提案されている手法を適用し、記述されたものが紹介されています。この場合も、ARSRG埋込の精度がその品質を裏付けている。

表8.COIL-100データベースにおける認識精度。

方法	精度
LLP+ARSRGemb	99.55%
LLP+BoW	51.71%
gdFil	32.61%
ブイエフエスア ール	91.60%
エーピーエム	99.11%
VEAM	99.44%
DTROD-AdaBoost （ドロッド・ア ダブースト	84.50%
RSW+ブースティ ング	89.20%
シーケンシャルパ ターン	89.80%
エルエフ	99.40%

[55]に関連する設定により ALOI データベースで得られた結果を表9にまとめる。特に、最初の100個の物体のみが採用されている。カラー画像はグレーレベルに変換され、学習用には各クラスの2番目の画像を、残りはテスト用に採用される。各クラス2枚ずつ、合計200枚の画像を得る。各反復の間に、各クラスについて、さらに1枚の学習用画像が添付される。表9は、中間結果では大きな差が出ないため、400枚のバッチを考慮した結果のみを示している。ベースラインの LLP（Logistic Label Propagation）[45] + BoW（Bag of Words）[46] と、線形判別分析のいくつかのバリエーション（ILDAaPCA、batchLDA、ILDAonK、ILDAonL）を適用して得られた結果[55]が報告されています。このように、LLP+ARSRGembは少ない学習セットで性能を発揮し、オーバーフィッティングの問題にもほとんど影響されないことが分かる。

表 9.ALOIデータベースにおける認識精度。

方法	200	400	800	1200	1600	2000	2400	2800	3200	3600
LLP+ARSRGem b	86.00%	90.00%	93.00%	96.00%	95.62%	96.00%	88.00%	81.89%	79.17%	79.78%
LLP+BoW	49.60%	55.00%	50.42%	50.13%	49.81%	48.88%	49.52%	49.65%	48.96%	49.10%
バッチエルデ イーエー	51.00%	52.00%	62.00%	62.00%	70.00%	71.00%	74.00%	75.00%	75.00%	77.00%
イルダペイカ	51.00%	42.00%	53.00%	48.00%	45.00%	50.00%	51.00%	49.00%	49.00%	50.00%
イルダオンク	42.00%	45.00%	53.00%	48.00%	45.00%	51.00%	51.00%	49.00%	49.00%	50.00%
イルダオンエ ル	51.00%	52.00%	61.00%	61.00%	65.00%	69.00%	71.00%	70.00%	71.00%	72.00%

さらに、局所的な情報を捉え、それらの間の空間的な関係を保存することで、物体認識の分野において強力な性能向上をもたらす。計算オーバーヘッドはグラフベースの表現構築のみであり、分類は非常に高速に実行できる。次元削減は非常に重要であり、それはグラフからベクトル空間への移行である。第一に、グラフのトポロジー、ノード間の関係、およびメイン構造のサブグラフに関するその他の関連情報を捉えること、第二に、グラフのサイズに関係なく処理の速度を下げる（グラフは一般に大きいので、良いアプローチは効率的でなければならない）、第三に、正しい次元を決定することである。長いベクトル表現は、より整理されたアプローチよりも、時間的・空間的に大きな複雑さを誘発しながらも、より多くの情報を保持する。要件に応じたバランスを見つけることが重要である。この場合、本手法は、まず、ARSRGの構造的特性をすべて保持することを試み、次に、目的空間に含まれるツールを活用することを試みる。埋め込み処理により、ベクトル空間は開始空間と比較して多くの処理ツールを含むため、より簡単に情報を表現し、分析することができる。

6.3. ARSRG フードバッグ

本節では、Bag of ARSRG Words (BoAW)と名付けられた[6]で以前に得られた結果をレビューする。

6.2節で述べたのと同じデータセットを採用し、さらに、以下に報告するデータセットも採用した。

- ・ カルテック101 [56]:101のカテゴリからなる天体画像集で、1カテゴリあたり約40～800枚の画像が収録されている。ほとんどのカテゴリで50枚程度の画像が収録されている。

ALOI、COIL-100、ETH-80の各データセットでは、背景に凹凸のあるCaltech 101データセットとは異なり、単純な背景に物体が表現されているため、分類段階の難易度は低くなっています。図5にオブジェクトの例を示します。

ディスカッション

以下の設定は[55]で報告されており、表9と同じものを採用した。分類段階ではLLP [45]、One-versus-All (OvA) パラダイムを30回実行、トレーニングセットとテストセットのシャッフル操作、150×150ピクセルサイズでの画像のスケーリングが採用されています。表10は、ALOIデータセットで、Bag of Visual Words (BoVW) [46]によって達成された精度結果と、線形判別分析のいくつかのバリエーション (ILDAaPCA、batchLDA、ILDAonK、ILDAonL) および[5] (ARSRGem b) を使用して得られた結果を報告している。

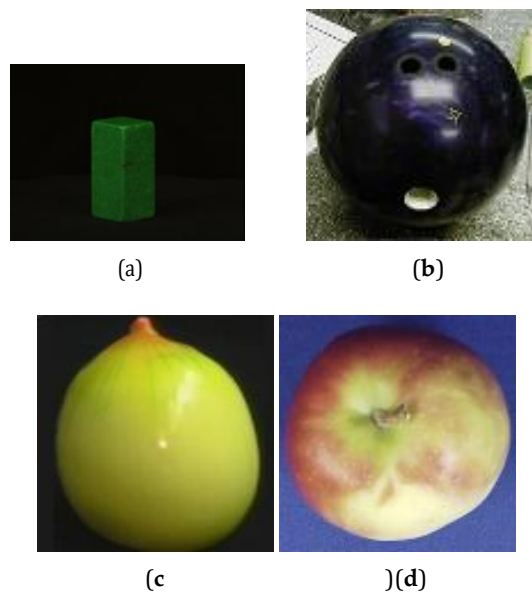


図 5.データセット画像：(a)ALOI, (b)Caltech 101, (c)COIL-100, (d)ETH-80。

表 10.ALOIデータセットでの結果。BoAW, Bag of ARSRG Words; BoVW, Bag of Visual Words.

方法	200	400	800	1200	1600	2000	2400	2800	3200	3600
BoAW	98.29%	92.83%	98.80%	96.80%	96.76%	98.15%	89.52%	82.65%	79.96%	79.88%
アールエ スアール ゲンブ	86.00%	90.00%	93.00%	96.00%	95.62%	96.00%	88.00%	81.89%	79.17%	79.78%
BoVW	49.60%	55.00%	50.42%	50.13%	49.81%	48.88%	49.52%	49.65%	48.96%	49.10%
バッチエ ルディー エー	51.00%	52.00%	62.00%	62.00%	70.00%	71.00%	74.00%	75.00%	75.00%	77.00%
イルダペ イカ	51.00%	42.00%	53.00%	48.00%	45.00%	50.00%	51.00%	49.00%	49.00%	50.00%
イルダオ ンク	42.00%	45.00%	53.00%	48.00%	45.00%	51.00%	51.00%	49.00%	49.00%	50.00%
イルダオ ンエル	51.00%	52.00%	61.00%	61.00%	65.00%	69.00%	71.00%	70.00%	71.00%	72.00%

その結果、物体認識タスクに適応することができるBoAWが最高の性能を発揮した。実際、主な貢献は局所情報と空間情報の組み合わせであり、画像表現とマッチングの段階を改善するものである。

Caltech 101データセットにおいて、特定の画像カテゴリ（ボーリング、ケーキ、電卓、大砲、CD、チェス盤、ジョイスティック、スケートボード、スプーン、傘）に対して、ピラミッド表現に基づくBoVW [46] と比較して得られた結果を表11にまとめている。訓練セットとテストセットをそれぞれ60/40%に分割した場合、最も良い平均精度が得られる。

表 11.Caltech 101 データセットでの結果。

メソッド	アキュラシー
BoAW	74.00
BoVW	83.00

画像が非一様な背景で構成されている場合の性能は異なることに容易に気がつくことができる。画像表現を歪め、結果として分類段階に影響を与えるこの側面から、BoVW は BoAW よりも強力である。可能な解決策としては、前処理の段階で、情報量の少ない背景を分離するためのセグメンテーションステップを設け、表現すべきオブジェクトのみを処理することが考えられる。この抜け道は、背景の除去が困難な場合があるため、必ずしも効率的ではありません。表12は、表8と同じ設定に基づくCOIL-100データセットの平均精度を示しています。したがって、

BoVW や [44,50] の解法 (VFSR) や [47] で提案されたアプローチ (gdFil) を適用して得られた結果との関連性を確認することができる。

[48] (APGM)、[49] (VEAM)、[51] (DTROD-AdaBoost)、[52] (RSW+boosting)、[53] (sequential patterns)、[54] (LAF)、[5] (ARSRGemb) で使用されています。さらに、この実験により、BoAWが最も良い性能を発揮することが確認された。

表 12.COIL-100データセットでの結果。

方法	精度
BoAW	99.77%
アールエスアール ゲンブ	99.55%
BoVW	51.71%
gdFil	32.61%
ブイエフエスア ール	91.60%
エービーエム	99.11%
VEAM	99.44%
DTROD-AdaBoost	84.50%
RSW+ブースティ ング	89.20%
シーケンシャルパ ターン	89.80%
エルエフ	99.40%

表 13 は、ETH-80 データセットと表 7 に関連するセットアップの結果である。BoVWに加えて、精度は[5](ARSRGemb), [47](gdFil), [48](APGM), [49](VEAM) で提案されたソリューションを採用して、[44]で達成したテストに関連しています。このように、視点変更ケースは、競合他社と比較してBoAWの性能に影響を与えないことがわかる。

このように、ARSRGは背景が一様であるような特定のタイプの画像に適している。具体的には、表現したいオブジェクトを前景とし、背景は無関係な情報を表していると考えることができる。画像の背景を走査しても特徴点は検出されないか、あるいは数点しか検出されない。背景は、通常、オブジェクトの符号化に有用な特徴点を含まないため、フィルタリングのステップを経て、前景から分離される。この方法は、Caltech 101データセットを除いて常に有効であり、性能に示されるように、場合によっては失敗する。

表 13.ETH-80データセットでの結果。

方法	精度
BoAW	89.29%
アールエ スアール ゲンブ	89.26%
BoVW	58.83%
gdFil	47.59%
エービー エム	84.39%
VEAM	82.68%

6.4. カーネルグラフ埋め込み

このセクションでは、以前[7]で得られた結果をレビューします。低、中、高バランスの画像分類問題において、標準的なOvA設定よりもサポートベクターマシン (SVM) と非対称カーネルスケーリング (AKS) [57] による分類性能が、芸術絵画分類アプリケーション [58] でテストされています。採用したデータセットは、セクション6.1で説明したものと同じである。表14と表15は、分類問題の設定を示しています。最後の列には、式 (5) により計算された、多数派クラスと少数派クラスに属する画像の割合である不均衡率 (IR) が含まれていることに注目されたい。

$$IR = \frac{\%maj}{分} \quad (5)$$

表 14.26]のデータセットに対するOvA(One-versus-All)構成。

問題点	分類の問題	(%min, %maj)	IR
1	アルテミシア vs. すべて	(3.00, 97.00)	32.33
2	バテシバ vs. オール	(3.00, 97.00)	32.33
3	ダナエ vs. オール	(12.00, 88.00)	7.33
4	Doctor_Nicolaes vs. all	(3.00, 97.00)	32.33
5	HollyFamily vs. all	(2.00, 98.00)	49.00
6	PortraitOfMariaTrip vs. All	(3.00, 97.00)	32.33
7	PortraitOfSaskia vs. All	(1.00, 99.00)	99.00
8	レンブラントXPortrai vs. all	(2.00, 98.00)	49.00
9	SaskiaAsFlora vs. all	(3.00, 97.00)	32.33
10	SelfportraitAsStPaul vs. All	(8.00, 92.00)	11.50
11	TheJewishBride vs. all.	(4.00, 96.00)	24.00
12	TheNightWatch vs. all	(9.00, 91.00)	10.11
13	予言者エレミヤとすべての人の比較	(7.00, 93.00)	13.28
14	TheReturnOfTheProdigalSon vs. all (放蕩 息子の帰還)	(9.00, 91.00)	10.11
15	TheSyndicsoftheClothmakersGuild vs. all.	(5.00, 95.00)	19.00
16	その他 vs すべて	(26.00, 74.00)	2.84

表 15.4]のデータセットに対するOvA構成。

問題点	分類の問題	(%min, %maj)	IR
1	クラス4 vs. すべて	(1.00, 9.00)	9.00
2	クラス7 vs. オール	(1.00, 9.00)	9.00
3	クラス8 vs. すべて	(1.00, 9.00)	9.00
4	クラス13 vs. オール	(1.00, 9.00)	9.00
5	クラス15 vs. オール	(1.00, 9.00)	9.00
6	クラス19 vs. オール	(1.00, 9.00)	9.00
7	クラス21 vs. オール	(1.00, 9.00)	9.00
8	クラス27 vs. オール	(1.00, 9.00)	9.00
9	クラス30 vs. オール	(1.00, 9.00)	9.00
10	クラス33 vs. オール	(1.00, 9.00)	9.00

ディスカッション

AKSと標準的なSVMを修正F値[59]の観点から比較した。図6より、AKSは標準的なSVMを上回っており、注目すべき性能に到達するためには、微調整が必要であることがわかる。また、図7では、SVMに対して1つのピークを超えるだけの性能である。さらに、OvA分類問題において、C4.5 [60], RIPPER [61], L2 loss SVM [62], L2 regularized logistic regression [63], Ripple-Down Rule learner (RDR) [64] との比較も行っている。不均衡率によってもたらされる歪みのため、データセットに関連する結果は異なっている。26]のデータセットでは、低、中、高の割合で構成されており、クラス不均衡問題の全ケースをカバーしているため、ロバストなテストフェーズに最適である。一方、[4]のデータセットでは、すべての構成でインバランス率が同じである。結果は、表16に[26]のデータセットについて、表17に[4]のデータセットについて報告されています。その結果、競合他社よりも著しく高い性能を示していることがわかる。特に、AKSの主な改善点は、関連性フィードバック評価においてより大きなウェイトを持つ少数派のクラス、positiveに属するパターンの分類精度に関するものである。実際、これらのパターンは、多数派のクラスである陰性に属するパターンと比べて分類が困難である。この結果は、2つの側面から、高いレベルの正しい分類に到達している。1つ目は、ベクトルベースの画像表現であるKGEARSRGを採用したことである。グラフカーネルは、特徴ベクトルの表現において、グラフの高い表現力と柔軟性の間のギャップを埋めることを目的としている。KGEARSRGは、分類のためのデータ処理を行うために、固定次元のベクトル空間画像表現を提供する。二つ目は、分類段階でのAKS法に関するものである。これは、以下のような分類問題に、より効率的に対処できる本質的な能力を持っている。

は極めてアンバランスである。つまり、AKS分類器は、多数派クラスに比べて少数派クラスから発生するパターンを正しく認識する能力を保持しているのです。

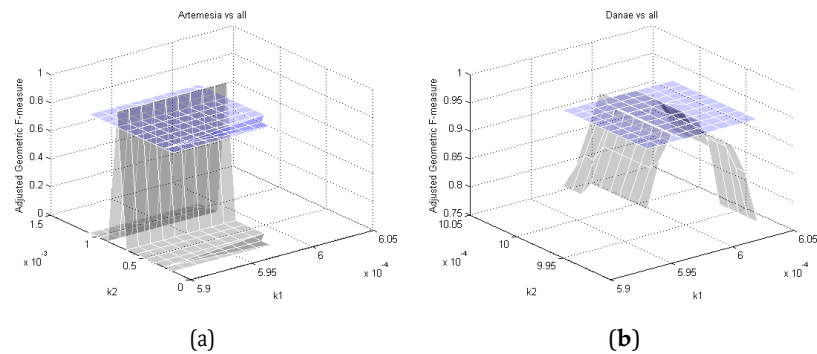


図 6.パラメータ選択1. x軸とy軸は2つの手法のパラメータの値を表し、z軸には[26]のデータセットのOvA構成のうち、**(a)** Artemisia vs. all と **(b)** Danae vs. all に対するAdjusted F-measure (AGF) がプロットされています。灰色と青色の面は、それぞれ非対称カーネルスケーリング (AKS) とSVM分類器による結果を表しています。

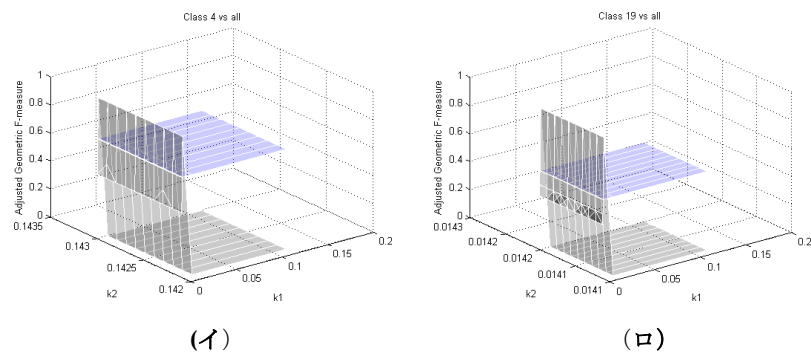


図 7.パラメータ選択 2.x軸とy軸は2つの手法のパラメータの値を表し、z軸は[4]のデータセットにおけるOvA構成のうち2つのAGFをプロットしたものである。

(a)クラス4対全体、**(b)**クラス19対全体。灰色と青色の面は、それぞれAKSとSVM分類器による結果である。

表 16.26]のデータセットと表 14 のデータセットで比較した結果。RDR, Ripple-Down Rule learner.

問題点	AGF					
	エーケー エス	C4.5	リッパ	L2-L SVM	L2 RLR	アール ディー アール
1	0.9414	0.5614	0.8234	0.6500	0.5456	0.8987
2	0.9356	0.8256	0.6600	0.8356	0.8078	0.7245
3	0.9678	0.8462	0.8651	0.4909	0.6123	0.7654
4	0.9746	0.8083	0.6600	0.4790	0.4104	0.6693
5	0.9654	0.7129	0.9861	0.8456	0.4432	0.6134
6	0.9342	0.5714	0.9525	0.8434	0.9525	0.5554
7	0.9567	0.6151	0.7423	0.5357	0.4799	0.6151
8	0.8345	0.4123	0.3563	0.7431	0.5124	0.7124
9	0.9435	0.9456	0.9456	0.8345	0.6600	0.6600
10	0.8456	0.4839	0.5345	0.4123	0.4009	0.5456
11	0.9457	0.9167	0.9088	0.9220	0.8666	0.9132
12	0.6028	0.5875	0.5239	0.4124	0.4934	0.5234
13	0.8847	0.7357	0.6836	0.7436	0.7013	0.5712
14	0.9376	0.9376	0.8562	0.8945	0.8722	0.8320
15	0.9765	0.8630	0.8897	0.8225	0.7440	0.8630
16	0.7142	0.5833	0.3893	0.4323	0.5455	0.5111

表 17.4]のデータセットと表15のデータセットでの比較結果。

問題点	AGF					
	エーケー ーエス	C4.5	リッパ	L2-L SVM	L2 RLR	アール ディー アール
1	0.9822	0.6967	0.5122	0.4232	0.4322	0.6121
2	0.9143	0.5132	0.4323	0.4121	0.4212	0.5323
3	0.9641	0.4121	0.4211	0.4213	0.3221	0.4323
4	0.9454	0.4332	0.1888	0.4583	0.3810	0.3810
5	0.9554	0.3810	0.2575	0.5595	0.3162	0.6967
6	0.9624	0.3001	0.1888	0.1312	0.3456	0.3121
7	0.9344	0.3810	0.5566	0.4122	0.4455	0.2234
8	0.9225	0.4333	0.1112	0.2575	0.1888	0.1888
9	0.9443	0.6322	0.1888	0.1888	0.6122	0.6641
10	0.9653	0.1897	0.5234	0.6956	0.1888	0.1121

7.結論

本論文では、画像表現のための構造であるAttributed Relational SIFT-based Regions Graph (ARSRG) を、新しい側面の記述と分析を通して探求している。先行研究から出発し、徹底的な研究を行い、ARSRGの構造設計を明確にし、深めるために、理論的な概念を導入している。ARSRGが、グラフマッチング、グラフ埋め込み、グラフ単語バッグ、カーネルグラフ埋め込みなどの異分野で、物体認識や絵画の検索・分類などの応用にどのように採用されるかが示される。実験結果は、様々なデータセットにおいて、最先端の競合技術よりも優れた性能を発揮することを十分に示している。今後の展開としては、さらなる応用分野の開拓、性能比較のための追加アルゴリズム（主にグラフマッチング）の導入、ARSRGに含めるべき画像特徴の充実などが考えられる。

資金提供この研究は、外部からの資金援助を受けていない。

謝辞この作品は、Alfredo Petrosinoに捧げます。彼とともに、私はコンピュータサイエンスの分野での最初の一步を踏み出しました。一緒に過ごしたこの数年間で、私は目標を達成するための堅固さと、仕事に対する愛と情熱を学びました。一生、感謝し続けます。ありがとう、私の偉大な師よ。

利益相反著者は利益相反を宣言していない。

参考文献

1. Love, B.C.; Roudier, J.N.; Wisniewski, E.J. A structural account of global and local processing. 大域的処理と局所的処理の構造的説明. *Cogn. 心理学* **1999**, *38*, 291-316. [\[CrossRef\]](#) [\[PubMed\]](#) を参照してください。
2. Koffka, K. *Principles of Gestalt Psychology*; Routledge: Abingdon, UK, 2013.
3. Liu, Y.; Zhang, D.; Lu, G.; Ma, W.Y. A survey of content-based image retrieval with high-level semantics. *パターン認識 (Pattern Recognit)* **2007**, *40*, 262-282. [\[クロスレフ\]](#)。
4. Manzo, M.; Petrosino, A. Attributed relational sift-based regions graph for art painting retrieval. In Proceedings of the International Conference on Image Analysis and Processing, Naples, Italy, 9-13 September 2013; pp.833-842.
5. Manzo, M.; Pellino, S.; Petrosino, A.; Rozza, A. A novel graph embedding framework for object recognition. (物体認識のための新しいグラフ埋め込みフレームワーク)。In Proceedings of the European Conference on Computer Vision, Zürich, Switzerland, 6-12 September 2014; pp.341-352.
6. Manzo, M.; Pellino, S. Bag of ARSRG Words (BoAW). *Mach. Learn. Knowl. Extr.* **2019**, *1*, 871-882. [\[クロスリーフ\]](#) [\(CrossRef\)](#)
7. Manzo, M. KGEARSRG: Kernel Graph Embedding on Attributed Relational SIFT-Based Regions Graph. カーネルグラフエンベッディングを用いた地域グラフの作成。マツハ。学ぶ。Knowl.Extr. **2019**, *1*, 962-973. [\[クロスリーフ\]](#) [\(CrossRef\)](#)
8. Lowe, D.G. スケール不変のキーポイントからの特徴的な画像特徴。 *Int. J. Comput. Vis.* **2004**, *60*, 91-110. [\[CrossRef\]](#) (英語)

9. SIFT記述子を用いた画像-特徴関連付けのための属性付きグラフマッチング. In *Structural, Syntactic, and Statistical Pattern Recognition*; Springer: Berlin, Germany, 2010; pp.254-263.
10. Sanroma, G.; Alquézar, R.; Serratos, F. A discrete labelling approach to attributed graph matching using SIFT features. (SIFT特徴を用いた属性付きグラフマッチングのための離散ラベリングアプローチ) .In *Proceedings of the 2010 20th International Conference on Pattern Recognition (ICPR)*, Istanbul, Turkey, 23-26 August 2010; pp.954-957.より。
11. Duchenne, O.; Joulin, A.; Ponce, J. A graph-matching kernel for object categorization. In *Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, 6-13 November 2011; pp.1792-1799.
12. Cho, M.; Lee, K.M. Progressive graph matching: Making a move of graphs via probabilistic voting. 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16-24 June 2012; pp.398-405.にて発表。
13. Matas, J.; Chum, O.; Urban, M.; Pajdla, T. Robust wide-baseline stereo from maximally stable extremal regions. *Image Vis. Comput.* **2004**, *22*, 761-767. [CrossRef] (英語)
14. Mikolajczyk, K.; Schmid, C. Scale & affine invariant interest point detectors (スケールとアフィン不変の関心点検出器) . *Int. J. Comput. Vis.* **2004**, *60*, 63-86.
15. Lee, J.; Cho, M.; Lee, K.M. Hyper-graph matching via reweighted random walks. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20-25 June 2011; pp.1633-1640.
16. Revaud, J.; Lavoué, G.; Ariki, Y.; Baskurt, A. Learning an efficient and robust graph matching procedure for specific object recognition (特定物体認識のための効率的で頑健なグラフマッチング手順の学習) . 2010年第20回パターン認識国際会議(ICPR), イスタンブール, トルコ, 2010年8月23-26日; pp.754-757.にて。
17. Romero, A.; Cazorla, M. 全方位画像を用いたトポロジカルスラム。特徴検出器とグラフマッチングを融合させる。 In *Advanced Concepts for Intelligent Vision Systems*; Springer: Berlin, Germany, 2010; pp.464-475.
18. Deng, Y.; Manjunath, B. Unsupervised segmentation of color-texture regions in images and video. 画像・映像における色・テクスチャ領域の教師なしセグメンテーション. *IEEE Trans. パターンアナライザー Mach. Intell.* **2001**, *23*, 800-810. [クロスリフ]
19. また、このような技術的な課題を解決するために、「また、このような場合にも、「曖昧さ」を解消するために、「曖昧さ」を解消するために、「曖昧さ」を解消するために、「曖昧さ」を解消するために、「曖昧さ」を解消するために、「曖昧さ」を解消する必要がある。構造的・統語的・統計的パターン認識」(Springer: Berlin, Germany, 2008)を開発し、その成果を発表した。
20. 堀 利明; 瀧口 智之; 有木 靖弘: グラフ構造表現による汎用的な物体認識. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25-30 March 2012; pp.1021-1024.にて発表。
21. Luo, M.; Qi, M. A New Method for Cartridge Case Image Mosaic. *J. Softw.* **2011**, *6*, 1305-1312. [CrossRef] (英語)
22. Trémeau, A.; Colantoni, P. Regions adjacency graph applied to color image segmentation. 地域隣接グラフのカラー画像セグメンテーションへの応用. *IEEE Trans. 画像処理* **2000**, *9*, 735-744. [CrossRef] (英語)
23. Mikolajczyk, K.; Schmid, C. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1615-1630. [CrossRef] [PubMed].
24. 劉 毅; 張 迪; 呂 庚; 馬 武毅. 知覚色を用いた領域ベースの画像検索. また、このような画像検索を実現するために、「画像検索技術」を開発した。
25. Sanromà Güell, G.; Alquézar Mancho, R.; Serratos Casanelles, F. Graph matching using SIFT descriptors- An application to pose recovery of a mobile robot. In *Proceedings of the Fifth International Conference on Computer Vision Theory and Applications*, Angers, France, 17-21 May 2010; pp.249-254.
26. Haladová, Z.; Šikudová, E. Limitations of SIFT/SURF based methods in classifications of fine art paintings. 「SIFT/SURFに基づく絵画の分類における限界」. *Comput. Graph. Geom.* **2010**, *12*, 40-50.
27. Chang, C.; Etezadi-Amoli, M.; Hewlett, M. A Day at the Museum. 2009. Available online: <http://www.stanford.edu/class/ee368/Project07/reports/ee368group06.pdf> (accessed on 6 August 2020).
28. また、このような場合にも、「適応的マルチメディア検索」の国際ワークショップを開催する。
29. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: 高速化されたロバスト特徴. In *Proceedings of the European Conference on Computer Vision*, Graz, Austria, 7-13 May 2006; pp.404-417.
30. このため、このような場合にも、SIFTやSURFに代わる効率的な手法として、ORBを用いることがで

31. Alahi, A.; Ortiz, R.; Vandergheynst, P. フリーク。高速な網膜のキーポイント。2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16-21 June 2012; pp.510-517にて発表。
32. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. Brief: 二値ロバスト独立素性。In Proceedings of the European Conference on Computer Vision, Heraklion, Crete, 5-11 September 2010; pp.778-792.
33. Cho, M.; Lee, J.; Lee, K.M. Reweighted random walks for graph matching (グラフマッチングのための再加重ランダムウォーク)。欧州コンピュータビジョン会議 (Heraklion, Crete, 5-11 September 2010) 発表論文; pp.492-505.
34. Duchenne, O.; Bach, F.; Kweon, I.S.; Ponce, J. A tensor-based algorithm for high-order graph matching. *IEEE Trans. パターンアナライザー-Mach.Intell.* **2011**, *33*, 2383-2395.[CrossRef] [PubMed].
35. Lux, M.; Chatzichristofis, S.A. Lire:Lucene 画像検索: 拡張可能な java cbir ライブラリ。In Proceedings of the 16th ACM International Conference on Multimedia, Vancouver, BC, Canada, 27-31 October 2008; pp.1085-1088.
36. MPEG-7規格の概要。 *IEEE Trans.Circuits Syst. ビデオ技術(Video Technol.)* **2001**, *11*, 688-695.[クロスリフ]
37. 田村秀男, 森修一, 山脇俊之: 視覚に対応するテクスチャ特徴. *IEEE Trans.Syst.Man Cybern.* **1978**, *8*, 460-473.[クロスリファレンス]
38. CEDD: Color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In Proceedings of the International Conference on Computer Vision Systems, Santorini, Greece, 12-15 May 2008; pp.312-322.
39. Chatzichristofis, S.A.; Boutalis, Y.S. Fcth:Fcth: Fuzzy color and texture histogram-a low level feature for accurate image retrieval. 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services, Klagenfurt, Austria, 7-9 May 2008; pp.191-196.
40. Huang, J.; Kumar, S.; Mitra, M.; Zhu, W.J.; Zabih, R. *Image Indexing Using Color Correlograms*; 1997; Volume 97, p. 762. Available online: <http://www.cs.cornell.edu/~rdz/Papers/Huang-CVPR97.pdf> (accessed on 6 August 2020).
41. Nayar, S.K.; Nene, S.A.; Murase, H. *Columbia Object Image Library (Coil 100)*; Technical Report No. CUCS-006-96; コロンビア大学: ニューヨーク、ニューヨーク州、アメリカ、1996年
42. Geusebroek, J.M.; Burghouts, G.J.; Smeulders, A.W. The Amsterdam library of object images(アムステルダムオブジェクト画像ライブラリー). *Int.J. Comput.Vis.* **2005**, *61*, 103-112.[CrossRef] (英語)
43. Leibe, B.; Schiele, B. Analyzing appearance and contour based methods for object categorization. In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 16-22 June 2003; Volume 2, pp.2-409.
44. Morales-González, A.; Acosta-Mendoza, N.; Gago-Alonso, A.; García-Reyes, E.B.; Medina-Pagola, J.E. 頻繁な近似サブグラフを用いたグラフベース画像分類のための新しい提案. *パターン・レコグニット*. **2014**, *47*, 169-177.[CrossRef] (英語)
45. ロジスティック・ラベルプロパゲーション. *Pattern Recognit.* **2012**, *33*, 580-588.[CrossRef] (英語)
46. Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: このような場合、「曖昧さ」を解消するために、「曖昧さ」を解消する方法として、「空間ピラミッドマッチング」がある。また、このような場合にも、「曖昧さ」の解消を図るため、「曖昧さ」の解消を図るべく、「曖昧さ」の解消を図るべく、「曖昧さ」の解消を図るべく、「曖昧さ」の解消を図るべく、「曖昧さ」の解消を図るべく、「曖昧さ」の解消を図るべく、「曖昧さ」を解消する。
47. Gago-Alonso, A.; Carrasco-Ochoa, J.A.; Medina-Pagola, J.E.; Martínez-Trinidad, J.F. Full duplicate candidate pruning for frequent connected subgraph mining (頻出連結部分グラフマイニングのための完全重複候補)。このような場合、「曖昧さ」を解消することが重要です。 *Comput.-Aided Eng.* **2010**, *17*, 211-225.[CrossRef] (英語)
48. Jia, Y.; Zhang, J.; Huan, J. An efficient graph-mining method for complicated and noisy data with real-world applications (複雑でノイズの多いデータに対する効率的なグラフマイニング手法)。 *Knowl.Inf.Syst.* **2011**, *28*, 423-447.[CrossRef] (英語)
49. Acosta-Mendoza, N.; Gago-Alonso, A.; Medina-Pagola, J.E. Frequent Approximate Subgraphs as features for graph-based image classification (グラフに基づく画像分類のための頻出近似部分グラフ)。を用いた画像分類。 **2012**, *27*, 381-392.[CrossRef] (英語)
50. Morales-González, A.; García-Reyes, E.B. 空間関係および不規則ピラミッドを用いて表現された視覚的特徴に基づく単純な物体認識。 *Multimed. ツールズ・アプル* **2013**, *63*, 875-897.[CrossRef].
51. Wang, Y.; Gong, S. Tensor Discriminant Analysis for View-based Object Recognition. 第18回パターン認識

52. Marée, R.; Geurts, P.; Piater, J.; Wehenkel, L. Decision Tree and random subwindows for object recognition (決定木とランダムサブウィンドウによる物体認識). *ICML Workshop on Machine Learning Techniques for Processing Multimedia Content (MLMM2005)*; University of Liege: リエージュ、ベルギー、2005年。
53. 森岡伸行: 順序パターンを用いた物体表現の学習.AI2008にて。 *Advances in Artificial Intelligence*; Springer: Berlin, Germany, 2008; pp.551-561.
54. Obdrzalek, S.; Matas, J. Distinguished Regions上のローカルアフィン・フレームを用いた物体認識.In *Proceedings of the British Machine Vision Conference 2002*, Cardiff, UK, 2-5 September 2002; Volume 2, pp.113-122.
55. Uray, M.; Skocaj, D.; Roth, P.M.; Bischof, H.; Leonardis, A. Reconstructive and Discriminative Approachesを組み合わせたインクリメンタルLDA学習。英国マシンビジョン Conference 2007, Warwick, UK, 10-13 September 2007; pp.1-10.
56. Li, F.F.; Fergus, R.; Perona, P. 少ない学習例から生成的な視覚モデルを学習する。101のオブジェクトカテゴリでテストされたインクリメンタルベイズアプローチ。 *Comput.Vis.Image Underst.*2007, 106, 59-70.
57. Maratea, A.; Petrosino, A. 非均衡データ分類のための非対称カーネルスケーリング。 *国際的なワークショップ「ファジールジックとその応用」*; Springer: Berlin, Germany, 2011; pp.196-203.
58. Ćiljak, M.; Mikuš, B.; Jež, K.; Hadjić, S. ジャンルによる美術絵画の分類。 In *Proceedings of the 2011 34th International Convention MIPRO*, Opatija, Croatia, 23-27 May 2011; pp.1634-1639.
59. Maratea, A.; Petrosino, A.; Manzo, M. Adjusted F-measure and kernel scaling for imbalanced data learning (不均衡データ学習のための調整済みFメジャーとカーネルスケーリング)。 *Inf.Sci.* 2014, 257, 331-341.[CrossRef] (英語)
60. クインラン, J.R. *C4.5: Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2014.
61. Cohen, W.W. Fast Effective Rule Induction.In *Machine Learning Proceedings 1995*; Elsevier: Amsterdam, The Netherlands, 1995; pp.115-123.
62. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, PA, USA, 27-29 July 1992; pp.144-152.
63. Fan, R.E.; Chang, K.W.; Hsieh, C.J.; Wang, X.R.; Lin, C.J. LIBLINEAR: A library for large linear classification (大規模線形分類のためのライブラリ)。 *J.Mach.Learn.Res.* 2008, 9, 1871-1874.
64. Dazeley, R.; Warner, P.; Johnson, S.; Vamplew, P. The Ballarat incremental knowledge engine.このような場合、「知識獲得ワークショップ」を開催する。



著者による©2020。ライセンシー: MDPI, Basel, Switzerland.この記事は、クリエイティブ・コモンズ表示 (CC BY) ライセンス (<http://creativecommons.org/licenses/by/4.0/>) の条件に基づいて配布されるオープンアクセス記事です。