

色情報と形状情報を用いた二次元画像の物体認識

An Object Recognition from Pictures using Information of Color and Form

阿部 敬由*¹ 奥村 紀之*²
 Noriyuki Abe Noriyuki Okumura

*¹長野工業高等専門学校
 Nagano National College of Tecnology

This research aims to construct an object recognition system from pictures using information of color and form. In this research, recognition targets are limited to vegetables and fruits. This system determines object using Knowledge-base which is defined by related words with objects. In addition, this paper suggests which the proposal system can be extended easily by discerning color and form of objects.

1. はじめに

近年の情報技術の発展に伴い、人工知能という研究が進められてきた。この人工知能をデバイスに搭載した、人間や動物に近い様々な動作を可能とするロボットが登場してきている。人工知能を搭載したロボットはほんの一部であったり、完全ではないが人間のある器官と同様の作業ができる。この人工知能という人の思考、動作を再現したものを利用し、福祉に役立てることができないかと考えている。そこで、本研究では人間の思考と関連付けた物体認識システムの構築を提案する。

本研究は、色や形状の情報を用いて、画像の中に含まれている物を判定する物体認識を行うことを目標としている。いきなりあらゆるものを認識するのは難しいので、今回は認識対象を特に野菜と果物としている。本システムは色と形状の特徴を「赤」、「丸い」等といった言語情報として抽出し、言語と物体を対応付ける知識ベース*¹により物体を特定するシステムである。そのため、物体の色と形状が分かれば、容易に認識対象を拡張できることを示している。

2. システムの概要

本研究では、写真等の情景画像*²から、色情報と形状特徴を取得し、その情報を知識ベースと照らし合わせて物体の識別を行う物体認識システムを提案する。図1にその物体認識のシステム流れを示す。

図1のように、情景画像に対し物体抽出*³を行い、それぞれの物体の色情報と形状情報を取得する。その情報を元に知識ベースを検索することによって、最終的にその情景画像には何が写っているのかを答える。色情報とは物体が持つ主要な色情報のことで、「赤」や「青」等と

連絡先: 奥村 紀之, 長野工業高等専門学校電子情報工学科, 026-295-7133, noriyuki_okumura@ei.nagano-nct.ac.jp

*¹ 人間の様々な物や事に対する知識が格納されているデータベース。Knowledge Base(KB)ともいう。

*² デジタルカメラ等で光景を撮影したもの。

*³ 情景画像と認識対象を分別すること。

いった言葉の情報として取得する。形状情報とは物体の概形情報のことで、「円」や「四角」といった言葉の情報として取得する。

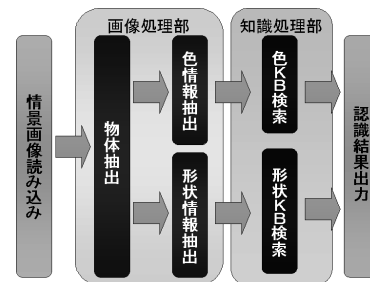


図1: 物体認識システムの流れ

3. 知識ベース構築のための調査

本システムに使用する知識ベースは、物体等に対しての人間の持つイメージの情報を格納している。本研究では、この知識ベースを作成するために、主に15~20歳の学生200名を対象にアンケートを行った。表1は、採ったアンケートの内容の1部を示したものである。

表1: アンケートの1部(例)

物体	形	色1	色2	備考
バナナ	三日月	黄	茶	腐ると茶色
林檎	丸	赤	茶	へたが茶色
蜜柑	丸	オレンジ	緑	へたが緑色
じゃが芋	丸	茶		

調査の内容は、設定した認識対象が、どのような形や色に見えるかを調査したものである。この目的は、人間が物体から連想する言葉を調査し、それを元に知識ベースを作成することにより、本システムと人間の思考の互換

性を高めることである。以後アンケートとはこのことを指すものとする。

4. 物体抽出

4.1 領域分割

本システムでは、扱う画像の条件が複雑なため、HSB 色空間における色相を用いて領域分割を行っている。しかし、色相で領域分割を行う場合、2つの問題点がある。1つは、画像では光の影響によって、同じ色でも明るい色と暗い色に変化してしまうことである。もう1つは、ディスプレイに出力されている画像は同じ色に見えても、内部では色が微妙に異なっているということである。色相を用いて領域分割を行う際にはこの2つの問題点を解決しなければならない。解決策としては、前者は HSB 色空間を用いることによって解決している。後者は色の量子化を行うことによって解決している。



図 2: 光の影響を軽減する前の画像



図 3: 色の量子化を行った画像

4.2 ラベリング

領域分割では、色相ごとに画像の領域を分割した。しかしこの状態では、どこからどこまでが1つの領域なのかコンピュータは識別できない。そこで、分割した領域に管理番号を与え、区別する必要がある。この管理番号を与える作業をラベリングという。本システムでは上下左右に隣接した同じ色の画素の集まりを1つの領域と考え、各領域ごとに管理番号を付けている。

5. 色情報抽出

5.1 基本色

物体を認識するにあたって、色は重要な要素である。コンピュータ上では色情報を数値で扱っている。しかし大抵の人間はその数値を見てもそれがどのような色なのかは判らない。人間の知覚とシステムに互換性を持たせるためには、色を数値ではなく、言葉として扱うのが適切である。また、最終的に野菜、果物だけではなく、あらゆるものの物体認識に対応するために、ある単語と単語の関連度が登録されており、膨大な数の単語が登録されている概念ベース [1] の利用を想定すると、色を名詞や形容詞で表す必要がある。本システムでは言葉として色の情報を取得するために、人間の感覚に対応するコンピュータ上の色を基本色として定義している。

5.2 基本色の定義

基本色は JIS で定義されている基本色 [2] を基に作成している。アンケートによる調査の結果により、JIS で定義された基本色のうち、全く登場しなかった「黄緑」、「青緑」、「青紫」、「赤紫」を削除している。また、頻繁に登場した「茶」、「ピンク」の2色を追加し、「白」、「黒」、「灰」、「赤」、「青」、「緑」、「黄」、「紫」、「オレンジ」、「ピンク」、「茶色」の11色となっている。基本色のうち、物体の中で最も多く使われている色を、色情報として取得する。

5.3 色の量子化

本研究では物体の中で最も多く使われている色を色情報として抽出する。コンピュータでは画像は RGB の組み合わせで表現されている。そのため RGB の約 1700 万色から基本色 11 色に変換を行う必要がある。そこで、色の量子化という考えが必要になる。色の量子化を実現するために、 $L^*a^*b^*$ 色空間 [3] を用いている。RGB 色空間から $L^*a^*b^*$ 色空間に変換した後、 $L^*a^*b^*$ 色空間上に定義されてある 11 の基本色との色差 ΔE を算出し、それぞれを比較する。この色差 ΔE が最も小さいものが、人間が心理的に近いと感じる色であるため、その基本色に変換する。

6. 形状情報抽出

6.1 テンプレートマッチング

形状の認識にはあらかじめテンプレートを用意しておき、対象物に一番近いものを選ぶテンプレートマッチングが一般的である [4]。この手法は、認識対象の形があらかじめ限定されている場合には、非常に有効である。しかし、本研究の場合は、認識対象の形が限定されていないため、多数のテンプレートが必要となり、この手法の適用は現実的ではない。そこで、本研究では物体からいくつかの形状の特徴を取得し、その特徴により、テンプレートを変形させてマッチングを行う可変型テンプレートマッチング方式を採用する。テンプレートには円形と多角形の2種類を用意する。円形テンプレートと多角形テンプレートで、それぞれでマッチングを行い、一致度が高い形状を採用する。一致度とは、テンプレートと物体において、画素同士がどの位の割合で一致しているかを計算するもので、式 (1) で求められるとする。

$$\text{一致度} = \frac{\text{テンプレートと物体の画素の積}}{\text{テンプレートと物体の画素の和}} \quad (1)$$

この操作により、物体の外形が「円形」か「多角形」という大まかな形状特徴を抽出することが出来る。

6.2 テンプレートの生成

まず、円形テンプレートの生成について説明する。物体の重心や大きさ等の情報がなければテンプレートマッチングに意味のあるテンプレートを作成することができない。そこで、次の方法を用いて、円形テンプレートを作成する (図 4 参照)

- 操作 1. 上端, 下端, 左端, 右端の 4 つの点を取得する.
- 操作 2. 4 つの点で最も距離が長い 2 点を結んで長軸とする.
- 操作 3. 長軸に垂直な線分を認識対象と重なる領域で生成し, 最も長くなるものを短軸とする.
- 操作 4. 生成した短軸の中心を直軸の中心に移動させる.
- 操作 5. 長軸と短軸を元に円形テンプレートを作成する.

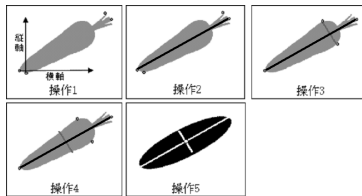


図 4: 円形テンプレート生成手順

- 多角形テンプレートの生成手順は次のとおりである.
- 操作 1. 上下左右の端の 4 点を取る.
 - 操作 2. 操作 1 で取得した 4 つの点を繋いで多角形を作成する.

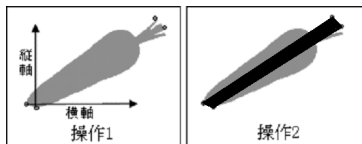


図 5: 多角形テンプレート生成手順

6.3 形状情報取得の流れ

生成した円形テンプレートと多角形テンプレートの一貫性を比較し, 高いほうをおおよその形状として取得する. ただし, 両方のテンプレートとも一貫性が極端に低いときは, 信頼性に欠けるため, 形状情報は不明となる. おおよその形状が決まった後, 更に細かい条件により, 最終的な形状が決定する. 最終的な形状の決定の条件については表 2 に示す.

表 2: 詳細な形状特徴のパラメータ設定

大まかな形状	形状	判定条件
円形	円	短軸 : 長軸 = 1 : 1.1 以下
円形	楕円	短軸 : 長軸 = 1 : 2 以下
円形	細長い	短軸 : 長軸 = 1 : 2 以上
多角形	細長い	短軸 : 長軸 = 1 : 2 以上
多角形	三角	4 点のうち 2 点の距離が極端に短いとき
多角形	四角	「細長い」, 「三角」に該当しないとき

7. 知識ベース検索

7.1 知識ベースの構築

本システムで使用する色知識ベース, 形状知識ベースをアンケートを基に作成している. 作成した知識ベース

を表 3, 表 4 に示す.

表 3: 色知識ベース

基本色	物体	基本色	物体
白		黒	
灰		赤	林檎, 桜ん坊, 苺, トマト
橙	人参, 柿, 蜜柑	黄	バナナ, レモン
緑	胡瓜, メロン, 西瓜	青	
紫	葡萄	茶	じゃが芋
ピンク	桃		

表 4: 形状知識ベース

形状	物体
円	蜜柑, 林檎, 桜ん坊, 葡萄, トマト, メロン, 西瓜
楕円	バナナ, じゃが芋, レモン, 柿
三角	人参, 苺
四角	
細長い	胡瓜

7.2 知識ベース検索

色情報抽出, 形状情報抽出で得た情報を元に知識ベースを検索する. 取得した色情報を元に, その色と関連付けられている物体を探し出し, 同様に取得した形状情報を元に, その形状と関連付けられている物体を探し出す. 色と形状に関連付けられた物体を探し出したら, その中で一致する物体を探し出す. 一致する物体があったら最終的な認識結果として出力する. 例えば「黄」, 「楕円」という情報が与えられているとき, 黄色に関連する物体はバナナ, レモンとなる. 一方楕円に関連する物体はバナナ, じゃが芋, レモン, 柿となる. この場合, 両方の知識ベース検索で一致したバナナ, レモンが最終的な認識結果となる.

8. システムの評価

各処理とシステム全体の認識精度について評価した. 以降に評価した条件と, 評価の結果を示す.

8.1 物体抽出

評価に使用した画像は, 本研究の認識対象が写っている情景画像である. 評価方法は 20 枚の情景画像内に写っている全認識対象のうち, 物体抽出を行えたものがいくつあったかで認識精度を計算した. 物体抽出において, 認識対象が情景から抽出できているか, 認識対象 1 つ 1 つが区別できているかという判断は 5 人が目視で判断した. その内 4 人以上が正常に行えていると判断したものを成功, そうではなかったものを失敗としている. 結果, 認識精度は 95% となった.

8.2 色情報抽出

評価に使用した画像は, 図 6 のように, 物体抽出されている画像を使用した. 評価方法は各野菜, 果物に対し 20 枚の画像を使用した. 画像に対し色情報抽出を行い, 知識ベースに登録してある情報と一致したものを成功, なかったものを失敗とした. 例えば, 林檎の画像に対し色情報抽出を行った場合, 色情報が赤という結果になれば成功である. 表 5 に色情報抽出の認識精度を示す.

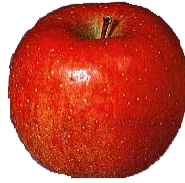


図 6: 情報抽出に使用した画像のサンプル

表 5: 各認識対象に対する色情報抽出の認識精度

認識対象	林檎	バナナ	人参	桜ん坊	胡瓜
認識精度	100	100	100	100	100
認識対象	葡萄	柿	レモン	メロン	蜜柑
認識精度	60	100	100	5	100
認識対象	桃	じゃが芋	莓	トマト	西瓜
認識精度	5	0	100	100	100

8.3 形状情報抽出

評価に使用した画像は図 6 のように、物体抽出されている画像を使用した。評価方法は各野菜、果物に対し 20 枚の画像を使用した。画像に対し形状情報抽出を行い、知識ベースに登録してある情報と一致したものを成功、しなかったものを失敗とした。例えば、林檎の画像に対し色情報抽出を行った場合、形状情報が円という結果になれば成功である。表 6 に形状情報抽出の認識精度を示す。

表 6: 各認識対象に対する形状情報抽出の認識精度

認識対象	林檎	バナナ	人参	桜ん坊	胡瓜
認識精度	100	70	85	95	95
認識対象	葡萄	柿	レモン	メロン	蜜柑
認識精度	0	95	100	100	90
認識対象	桃	じゃが芋	莓	トマト	西瓜
認識精度	100	100	55	85	100

8.4 知識ベース検索

色情報抽出、形状情報抽出にて知識ベースに登録してある色情報、形状情報が得られたものとして、各野菜、果物の情報に対し 1 回知識ベース検索を行い、評価した。例えば、林檎の場合、既に赤、円という情報があるものとして、その情報を元に知識ベース検索を行っている。出力結果の中に与えた情報に対する的確な回答が含まれていれば成功とし、回答の物体が複数あった場合でも、その中に的確な回答が含まれていれば成功としている。結果、認識精度は 100%となった。

8.5 システム全体

評価に使用した画像は、本研究の認識対象が写っている情景画像である。20 枚の情景画像内に写っている全認識対象のうち正常に物体認識を行えたものがいくつかあったので認識精度を求めた。回答の中に的確な回答が含まれていれば成功としている。システム全体の認識精度と各処理の認識精度を表 7 に示す。

表 7: 各認識対象に対する色情報抽出の認識精度

処理部名	物体抽出	色情報抽出	形状情報抽出
認識精度	95	78	85
処理部名	知識ベース検索	システム全体	
認識精度	100	71	

9. システムの考察

システム全体の認識精度は 71%となった。認識精度を下げている原因としては色情報抽出、形状情報抽出が挙げられる。その中でも特にいくつかの認識対象の認識精度が極端に低いことが原因である。そのため、全体の認識精度を向上させるには、こういった特定の認識対象の認識精度を上げる必要がある。

システム全体の認識精度から、本システムが導き出す情報と人間の連想する言葉、本システムの知識ベースと人間の知識の関連付けをすることができ、本システムの有用性を示せたと考える。

10. おわりに

本研究では、色情報と形状特徴を用いた物体認識を行う手法を提案した。認識精度は全体的に見れば 71%という精度だが、その内容は認識対象ごとに極端に精度が異なるものであった。今後の課題は、現時点でのシステムの短所を克服し、全体的に精度が高めることである。そうすれば認識対象に対してより精密な情報が得られ、いずれは野菜、果物以外の物体も認識できるようになると考える。

今回の研究により、情景画像からの物体抽出、画像から得られる情報と人間の連想する言語との関連性を持たせることの可能性について実証することができた。また、情報抽出によって確実な情報が得られるとすれば、本システムのために製作した知識ベースだけでなく、既存の概念ベースを用いることができるため、より物体認識の幅が広がると考える。

参考文献

- [1] 奥村紀之, 土屋誠司, 渡部広一, 河岡司 概念間の関連度計算のための大規模概念ベースの構築 自然言語処理. 2007 年 10 月, vol.14, No.15 pp.41-64
- [2] 日本規格協会 (2002) 『JIS ハンドブック 61 色彩』日本規格協会 836pp.
- [3] 篠原克幸 南敏 (1991) 『L*a*b*均等色空間におけるカラーエッジ抽出法』テレビジョン学会技術報告 Vol.15, No.29(19910523) pp. 19-24.
- [4] 酒井幸市 (1997) 『デジタル画像処理入門』コロナ社 167pp.
- [5] 重田紗綾 (2004) 『色系統と形状特徴による物体認識』同志社大学卒業論文.