

画像処理による計数・数詞教育システム

飯沢 奈緒 (指導教員：椎尾 一郎)

1 はじめに

数は私たちの生活の中にあふれており、日常生活を送る上で欠かせない。その反面極めて汎用性が高く抽象的なものであるため、子どもが言葉を学び数を理解するには多くの段階を経る必要がある。実際、子どもが数を覚える際に“唱える”ことから入り、音でしか認識してない時期があるとも言われている。数を1から言うことはできるが、それが実際の物の個数などと一致していないのである [1]。

一方、大人が外国語を学ぶ際にも数は必要不可欠なもので、最初に学ぶべきもののうちの一つとする人も多い。初めて学ぶ外国語においても、数は抽象的な概念であるため、覚えた通りの順番を考えなければ思い出せないという経験をしたことがある人もいないだろうか。

そこで本研究ではこのような数の理解と学習を対象にした学習システムを開発する。ここでは、仮想のエージェントが、日常目に見ている様々な物を数える様子をユーザに見せることで、数を学習する機会を提供する。このため、日常生活で物をカウントするシステムを提案し、そのための初期段階の実装を行った。

2 カウントシステム

子ども向け教育番組として有名なセサミストリートに、カウント伯爵 (Count von Count / The Count) ¹ というキャラクターがいる。彼は数が大好きで、同じ物が並んでいるとどこからともなく出てきて、数を数え上げる。本システムの目的は、このカウント伯爵を拡張現実世界に作ることである。すなわち、いつもユーザの近くにおいて、同じものを見つけては数え上げるシステムである。これをヘッドマウントディスプレイ (HMD) のようなウェアラブル機器に実装すれば、ユーザが数えようと意識しないときでも勝手に数え上げる仮想キャラクターを実現できる。これにより、学習のための時間を取ることなく、数に触れる機会を増やすことができ、数を覚えることができるであろう。また、数え上げる言語を外国語にすることで、外国語を学ぶ大人にも有用なシステムとなるであろう。

このように物と関連づけてカウントすることにより、単に数を唱えるときとは異なり、視覚的にも数に触れることができる。その結果、順序としての数だけではなく、一つ一つの独立した言葉として記憶に残りやすいと考えている。

3 実装

本研究ではカウントシステムの前段階として、Android スマートフォンを用いた静止画像カウントシステムの実装を行った。Android は汎用性が高く、Glassware などの HMD にも用いることができるため、今後ウェアラブル機器が日用化された際に、それに容易に移行することが可能である。また、静止画を対象と

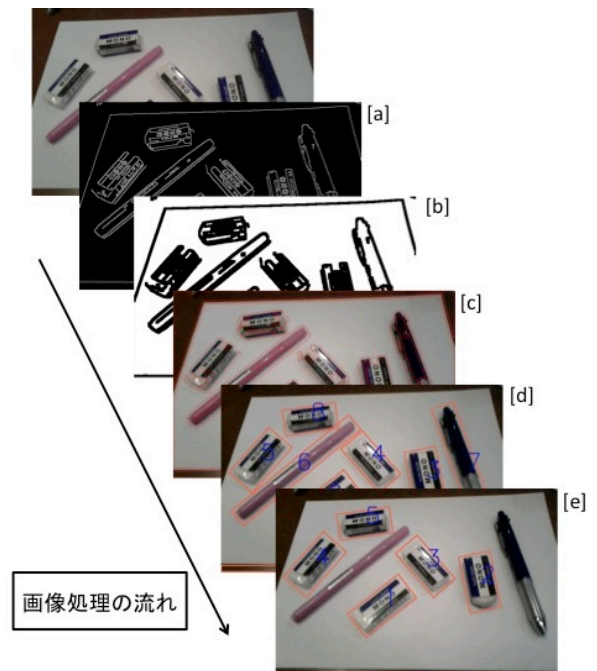


図 1: 主な画像処理

することで、処理が軽くなるメリットがある。さらに、この機能をデジタルスチルカメラに組み込むことで、物を数える機能をもったデジカメが実現可能になる。写真に写っている物体を自動的に数え上げる機能は、棚卸しなどの物品確認や計数作業など様々な場面で有用であると考えられる。

3.1 画像処理

本システムが、Android の内部で行う主な画像処理の流れを図 1 に示す。最初、スマートフォンには、カメラからの画像が表示されている。ユーザがこの画面を触ると、システムは静止画を撮影し、そのあと OpenCV を用いた画像処理により画像の中から同じ物を見つけ、音声で数え上げる。

以下、用いた主な画像処理について説明する。

- エッジ検出
入力画像からエッジを検出し、輪郭をはっきりさせる。(図 1 - [a])
- 二値化
エッジ抽出した画像を二値化する。その際に収縮処理を行い輪郭を強調する。(図 1 - [b])
- 輪郭抽出
二値化した画像から輪郭を抽出する。これにより輪郭線がポイントのリストとして抽出できる。(図 1 - [c])
- 輪郭を囲む矩形
抽出した輪郭を囲む回転矩形情報を取り出す。このとき画面に対して占領面積の大きい矩形、小さ

¹<http://sesame-street.jp/about/character.html>
http://en.wikipedia.org/wiki/Count_von_Count

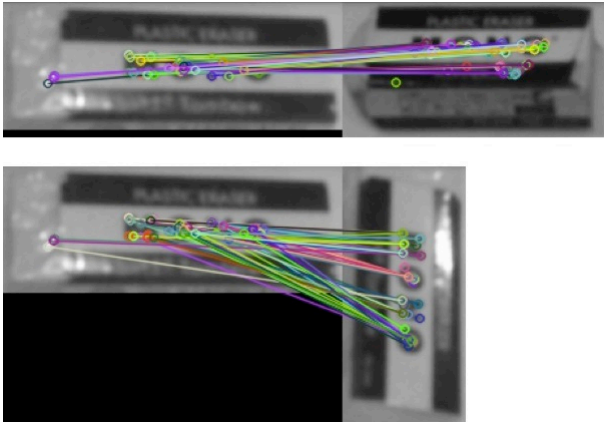


図 2: 特徴量マッチング

い矩形を取り除き、同時に重なっている矩形も取り除く。(図 1 - [d])

- 特徴量比較

画像から回転矩形を切り取り、その矩形ごとに検出器を用いて特徴点を抽出。その特徴点に対する特徴量を比較する。今回は処理の速い ORB 特徴量 [2] を用いた。マッチングの様子を図 2 に示す。

- 同一物判定

特徴量比較において閾値以下のマッチングの割合により、同一物を判定する。(図 1 - [e])

3.2 アプリケーション

本アプリケーションは、撮影写真から同一物を判定すると、その各々を画面上で矩形で囲みながら、同時に音声で数え上げる。その際には Android のテキスト読み上げ出力 (Text To Speech) を利用する。使用言語は多くの Android 端末で有効になっている Google テキスト読み上げ² に基づき、英語、フランス語、イタリア語、ドイツ語、韓国語、日本語を実装した。それぞれの言語を自由に選べるほか、ランダムにすることも可能である。

4 判定結果

図 3 に本システムによる成功画像と失敗画像を示す。現在は色による判別をしていないため、色違いのテープは同じ物としている。特徴量比較を用いているため、回転や縮小には対応しており、はっきりとした柄である消しゴムは三次元回転をしていても同一物判定が可能であった。一方で、小さいチップでは描かれているイラストが異なる物でも同一物判定をしてしまったり逆に同じ物でも違う物と判定したりしてしまうなど誤認識もみられた。この原因としては光による影や反射の違い、単色による特徴点の取り難さがあげられると考える。

5 関連研究

よく知られているカメラを用いた画像認識として顔認識 [3] があげられる。顔認識では大量の学習データに基づき顔の特徴検出に特化した認識手法となってい

²<https://play.google.com/store/apps/details?id=com.google.android.tts>



図 3: 左: 成功結果 右: 失敗結果

る。特定物体ではなくさまざまな物を同一物判定したため大量画像を扱う学習手法は難しいと考えた。

また、モバイル機器を使った物体認識として、モバイル物品管理システム [4] がある。これは初めての物を写真とテキストで登録し、次に同じ物を撮影した際に同じ物を判定して情報を出すというものである。

6 まとめと今後の課題

実世界の物を音と表示で数え上げるカウントシステムの提案を行い、その初期段階として Android スマートフォンのカメラを用いて静止画像で同一物判定によるカウントシステムを実装した。現段階では特徴量比較のみで同一物体を認識しているため、特徴点の少ない物体 (単一色) などは判定が難しくなっている。そのため、今後は同一物の判定のために特徴量比較以外のマッチングも行い、更に精度を高くしていきたい。加えて、物体の切り分け段階でも輪郭線の突然のゆがみなどを排除することで精度を上げることが可能ではないかと考える。また今回は、音声を音声合成という形で実装したが、数字のみの発声であり有限個であるため、音声ファイルを用意することも可能である。それにより、音声のエフェクトに伴った習得率なども調査していきたい。

以上をふまえ、処理をリアルタイムで行い、突然数え始めるようなシステムに繋げていくことを考える。

参考文献

[1] 平井 安久, 青山 陽一, 曾布川 卓也: 数の概念の捉え方, 数理解析研究所講究録 第 1828 巻 86-100 (2013)

[2] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski: ORB: an efficient alternative to SIFT or SURF, IEEE International Conference on Computer Vision (2011)

[3] P Viola, M J. Jones: Robust Real-Time Face Detection, International Journal of Computer Vision 57(2), 137-154 (2004)

[4] 望月 宏史, 柳井 啓司: 物体認識技術を用いたモバイル物品管理システム, DEIM Forum (2012)