

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 19 日現在

機関番号：51303

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330216

研究課題名(和文)統計的構造解析による画像認識用データの自動構築

研究課題名(英文)Automatic data generation for image recognition by statistical structure analysis

研究代表者

大町 方子(Omachi, Masako)

仙台高等専門学校・専攻科・准教授

研究者番号：90316448

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：画像処理や画像認識の分野では、大量データを用いた学習による手法が成果を上げている。しかし、大量で信頼できるデータを収集することは現実的には困難な場合が多い。本研究では、構造的特徴を反映させて画像データを生成するモデルを構築する手法を開発した。画像を構造的特徴に応じて自動的に分割し、構造ごとにモデル化する。そして、各構造のモデルを統合することで全体のモデルを得る。構築されたモデルを活用することにより画像認識の精度が向上することを確認した。

研究成果の概要(英文)：In the field of image processing and image recognition, a lot of techniques using big data have been proposed. However, it is difficult to collect enough reliable data in many cases. In this study, we developed a method for constructing a model that can generate image data by considering the structure. An image is divided into regions according to its structure. Then a model is constructed for each region. A whole model is constructed by combining all the models. Experimental results using the proposed model show its effectiveness.

研究分野：パターン認識

キーワード：画像、文章、音声認識等 パターン認識

1. 研究開始当初の背景

画像処理や画像認識の分野では、大量データを用いた学習による手法が多数研究され、成果を上げている。例えばインターネットから収集した大量の画像を用いることにより単純な手法でも高度な画像処理や画像認識が可能となることなどが実験的に示されている。これらの手法では、高精度な手法を実現するために大量で信頼できる(正確にラベル付された)データを収集することに労力が費やされている。しかし、一般的には大量で信頼できるデータを収集することは現実的には困難な場合が多く、何百万というデータがあることを前提とした手法は多くの場合そのまま利用することは難しい。

一方、スパースコーディングに代表されるように、現実世界のデータは少数の基底から構築することが可能という考え方もある。したがって、適切な手法を用いることにより、少数データから信頼できる大量のデータを構築することは不可能ではないと考えられる。

データを自動生成する試みはこれまでもあり、成果を上げている。しかし、既存の手法の多くはアピアランススペースであり、しかも対象に応じてデータ構築のための知識を人間が与える必要があった。事前知識を必要とせず、アピアランススペースではなく特徴ベクトルレベルでのデータ構築が可能となれば、応用範囲が広く、多くのアプリケーションにおいてアルゴリズムを変更することなく認識の高精度化が可能となる

2. 研究の目的

構造的特徴を反映させた画像データの自動構築を行なう手法を開発することを目的とする。アプローチとしては、特徴ベクトルベースの手法とアピアランススペースの手法をそれぞれ検討する。いずれも、できるだけ事前知識なしに実現することを目指す。

特徴ベクトルベースの手法としては、画像の特徴ベクトルを構造的特徴に応じて自動的に分割し、構造ごとにモデル化する。そして、各構造のモデルから生成されたパターンを統合することで全体のパターンを得る。

一方、アピアランススペースの手法としては、画像パターン全体を変形してデータを生成する方法と、パターンを分割して組み合わせることによりデータを生成する方法を検討する。

そして、構築されたデータを活用する画像認識手法を開発する。実際の画像データを用いて認識実験を行い、効果を確認する。

3. 研究の方法

まず手書き文字のように連続的に形状が変化するパターンを対象とし、既存のデータや独自に取得したデータを用いて可視化によりデータの分布形状を解析する。その結果を踏まえ、データを特徴ベクトルで表現して

特徴ベクトルの要素をクラスタリングし、各クラスタをモデル化する手法を開発する。これらのデータを統合することでデータの自動構築アルゴリズムを開発する。さらに、データ統合の過程を考慮することで、構築されたデータを活用する画像認識アルゴリズムを開発する。

一方、活字文字のように連続的な変化ではない様々なデザインが存在するパターンに対しては、上記とは異なるアプローチによりデータを構築する手法を開発する。具体的には、統計的な分布形状を考慮するのではなく、少数のサンプルをベースにした手法を検討する。

公開データベースおよび独自に収集したデータを用い、データ構築および構築されたデータを用いた認識実験を行い、提案手法の効果を確認する。

4. 研究成果

(1) 画像データのモデル化手法の開発

画像データの構造的特徴を反映させたデータの自動構築を行う手法の開発を行った。まず画像データをベクトルで表し、ベクトルの要素を相関に応じてクラスタリングすることで構造ごとにモデル化し、それらを統合してデータを生成する手法を開発した。ベクトルの要素をノード、要素間の相関をエッジ重みとするグラフを考え、グラフカットの手法を用いることで相関の大きい要素どうしが同じクラスタに属するように要素を分類することができる。この手法により、画像の各領域のうち連動して変形する部分とほぼ独立して変形する部分を分離することができる。連動して変形する部分ごとに統計的にモデル化を行うことで、妥当なモデルを構築することができる。

手書き文字データを用い、モデルを構築する実験を行った。特徴ベクトルとして、画素の輝度値を並べたものを用いた場合の結果を図1に示す。「2」の複数の文字パターンを入力として、上記の手法により2個のクラスタにクラスタリングし、クラスタごとの3つの典型的なパターンを組み合わせることで表示したものである。2個のクラスタのうち一方を黒画素、もう一方を灰色画素で表している。黒画素は「2」の上部をモデル化しており、T1~T3が傾きの異なる3つのパターンを表している。灰色画素は主に「2」の下部をモデル化しており、B1はループを含む形状、B3はそうでない形状、B2はこれらの中間の形状を表している。また、上部の一部が灰色画素となっているのは、この部分のみ下部との相関が大きいことを表している。すなわち、書き始めの部分と下部が連動して変形していることになる。このように、このモデル化手法は、人間の直感とは関係なく、統計的に妥当なデータのモデル化を自動的に行うことができる点に大きな特徴がある。

黒画素 灰色画素	T_1	T_2	T_3
B_1			
B_2			
B_3			

図1 「2」のモデル

本モデルは確率的なモデルであるため、モデルを直接使い、統計的パターン認識の枠組みで認識を実現することが可能である。手書き文字認識データベースを用いた認識実験により、精度が向上することを確認している。また、本モデルは画素の輝度値を並べたベクトルのみならず任意の特徴ベクトルに適用できる点も大きな特徴である。

(2) モーフィングを用いたデータ構築手法の開発

より直感的に構造情報を反映させるために、情景画像中の活字文字を対象とし、モーフィングを用いて文字データを生成する手法についても検討を行った。モーフィングは画像を滑らかに変形させる技術であるが、さまざまな活字フォントのように大きく異なる複数の文字に対して適用するには工夫が必要であり、サンプリングとスプライン曲線近似を利用した手法を開発した。文字部分と背景部分を構成するテクスチャ画像を生成する手法も開発し、文字パターンどうしのモーフィング結果と生成されたテクスチャ画像を融合することで、視覚的に妥当な文字パターンを生成する。

手法の有効性を確認するために、公開データベースの文字パターンおよびフォントパターンを用いてデータを構築する実験を行った。図2に結果の例を示す。2種類の「A」の文字から別の「A」の文字の概形を構築し、一方でテクスチャ情報から前景と背景を構築し、これらを統合することで新たな文字パターンを生成する様子を示している。

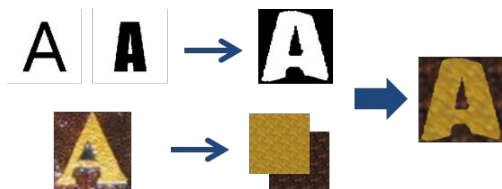


図2 モーフィングによるデータ生成

提案手法によって作成された人工パターンを用いて文字認識実験を行い、認識精度が向上することを確認した。

(3) パーツの分解と配置を利用したデータ構築手法の開発

モーフィングを用いた手法ではパターン全体を変形するのに対して、パターンを基本的なパーツに分解し、パーツを配置することでデータの生成を行う手法についても検討を行った。具体的には、活字文字を対象とし、個々のストロークをパーツとみなし、活字文字パターンからストロークを得る。ストロークを文字の概形に沿って配置することで、ストロークの特徴を保持した文字パターンを構築することができる。

この手法では、サンプル文字パターンとその骨格データを入力とする。骨格データの各ストロークに対応するサンプル文字パターン中のストロークを抽出し、パーツとする。正確な抽出のためにエネルギー最小化を用いた最適化を行う。図3に具体例を示す。複数のサンプルから抽出されたストロークを組み合わせることで、新しい文字パターンを生成する様子を示している。

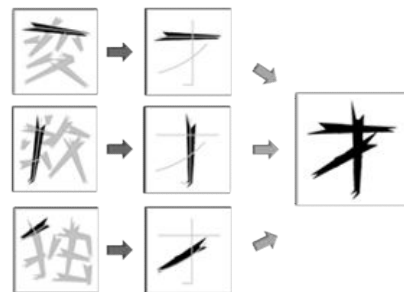


図3 パーツの分解によるデータ生成

提案手法によって作成された人工パターンを用いて文字認識実験を行い、認識精度が向上することを確認した。また、主観評価実験により、違和感のない文字パターンが生成できていることを確認した。

(4) 認識手法

データ構築手法の開発と並行して、複数のデータを活用する認識手法についても検討を行った。複数のデータから部分空間を構築し、部分空間どうしの類似性で認識を行う相互部分空間法をベースに、高速化、高精度化のための手法を開発した。第1固有ベクトルとその他の固有ベクトルの性質の違いに注目し、これらを分けて処理することにより高速化を達成している。

一方で、グラフを用いたより柔軟な変形に対応する認識手法についても検討を行った。統計的な考えを導入してグラフのモデルを構築するが、グラフの特性を利用することで統計的には扱いきれない大きな変形にも対応することが可能となる。

そして、実際のデータを用いた認識実験で、提案手法の効果を確認した。

5. 主な発表論文等

〔雑誌論文〕(計1件)

Hiroki Moizumi, Yoshihiro Sugaya, Masako Omachi, and Shinichiro Omachi, Traffic Light Detection Considering Color Saturation Using In-Vehicle Stereo Camera, Journal of Information Processing, 査読有, 24 巻, 2016 年, 349-357
DOI: 10.2197/ipsjjip.24.349

〔学会発表〕(計 13 件)

酒井利晃, 宮崎 智, 菅谷至寛, 大町真一郎, 非線形最小化によるグラフのモデルの構築と画像認識, 電子情報通信学会 2016 年総合大会, 2016 年 3 月 17 日, 九州大学(福岡市)

Airi Kitasato, Tomo Miyazaki, Yoshihiro Sugaya, and Shinichiro Omachi, Discrimination of Scomber Japonicus and Scomber Australasicus by Dorsal Fin Length and Fork Length, The 22nd Korea-Japan joint Workshop on Frontiers of Computer Vision, 2016 年 2 月 19 日, ひだホテルプラザ(高山市)

Shuto Shinbo, Tomo Miyazaki, Yoshihiro Sugaya, and Shinichiro Omachi, Accuracy Improvement of Character Recognition Using Generated Samples by Morphing, 平成 27 年度電気関係学会東北支部連合大会, 2015 年 8 月 28 日, 岩手県立大学(滝沢市)

Masako Omachi and Shinichiro Omachi, Construction of Models for Data Generation Considering Data Distribution, International Conference on Information Technology and Engineering, 2015 年 8 月 12 日, 北京(中国)

Shuhei Toba, Hirotaka Kudo, Tomo Miyazaki, Yoshihiro Sugaya and Shinichiro Omachi, Ultra-low Resolution Character Recognition System with Pruning Mutual Subspace Method, The IEEE International Conference on Consumer Electronics - Taiwan, 2015 年 6 月 7 日, 台北(台湾)

鳥羽修平, 工藤裕貴, 宮崎 智, 菅谷至寛, 大町真一郎, 動画像を用いた超低解像度文字認識, 電子情報通信学会パターン認識・メディア理解研究会, 2015 年 2 月 19 日, 東北大学(仙台市)

大町真一郎, 文字認識技術の変遷と最新動向, マシンビジョン研究会 第 2 回研究会, 2015 年 1 月 27 日, TKP ガーデンシティ仙台勾当台(仙台市)

Jian Wang, Hiroya Saito, Tomo Miyazaki, Yoshihiro Sugaya and Shinichiro Omachi, Automatic Synthesis of Character Pattern Using Patch Transform, 2014 International

Workshop on Electronics & Communications, 2014 年 10 月 27 日, 成都(中国)

Shuhei Toba, Hirotaka Kudo, Tomo Miyazaki, Yoshihiro Sugaya, and Shinichiro Omachi, Ultra-Low Resolution Character Recognition with Increased Training Data and Image Enhancement, 平成 26 年度電気関係学会東北支部連合大会, 2014 年 8 月 22 日, 山形大学工学部(米沢市)

Tatsunori Tsuchiya, Tomo Miyazaki, Yoshihiro Sugaya, and Shinichiro Omachi, Automatic Generation of Kanji Fonts from Sample Designs, 平成 26 年度電気関係学会東北支部連合大会, 2014 年 8 月 22 日, 山形大学工学部(米沢市)

大町真一郎, ビッグデータ時代の文字検出・文字認識技術, 精密工学会・画像応用技術専門委員会 サマーセミナー 2014, 2014 年 8 月 20 日, 岩松旅館(仙台市)

土屋達徳, 宮崎 智, 菅谷至寛, 大町真一郎, サンプルデザインからの漢字フォントの自動生成, 第 17 回画像の認識・理解シンポジウム, 2014 年 7 月 31 日, 岡山コンベンションセンター(岡山市)

Tatsunori Tsuchiya, Tomo Miyazaki, Yoshihiro Sugaya, and Shinichiro Omachi, Automatic Generation of Kanji Fonts from Small-Size Samples, 2013 International Workshop on Emerging ICT, 2013 年 10 月 29 日, 東北大学(仙台市)

〔図書〕(計 1 件)

白鳥則郎, 大町真一郎, 陳謙, 大町方子, 宮田高道, 長谷川為春, 早川吉彦, 加瀬澤正, 塩入諭, 共立出版, 画像処理, 2014 年, 232(1-8, 23-37, 175-189)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

大町 方子(OMACHI, Masako)

仙台高等専門学校・専攻科・准教授

研究者番号: 90316448

(2) 研究分担者

大町 真一郎(OMACHI, Shinichiro)

東北大学・大学院工学研究科・教授

研究者番号: 30250856