

料理写真の魅力度推定器を使った写真撮影補助システムの検討

大木 慧矢† マッキン ケネスジェームス† 永井 保夫†

†東京情報大学 総合情報学科

1. はじめに

近年スマートフォンの発展、普及により写真を撮るのが手軽となり写真を SNS などに上げることが容易になってきた。また、写真を撮る上で色合いなどは最近の HDR (High Dynamic Range) 機能や写真の編集アプリで調節することができるようになり、簡単に編集できてしまうため写真に対するクオリティーの高さが求められる。しかし、色が簡単に編集できるのに対し最適な構図を決めて上手く撮ることはなかなか難しい。この問題を解決するために、深層学習を利用し、最適な構図で撮影をする補助システムの開発を検討する。

佐藤ら[1]は、畳み込みニューラルネットワーク (CNN: Convolution Neural Network) である VGG16 が魅力度推定器に適していることを示した。

本研究ではこの魅力度推定器を参考にスマートフォンなどから撮影した料理の動画から最適な構図の写真を出力する撮影補助システムを検討する。

2. 提案手法

2.1 魅力度推定器

提案手法の魅力度推定器では、VGG16 および VGG19 に転移学習を行う。転移学習をするにあたって井手[2]らの作成したデータセット「NUFOOD 360×10」を用いる料理の写真を 244×244 にリサイズし入力し、画像に対する魅力度を教師データとして学習を行う。重みの学習を行う全結合層の活性化関数には ReLU を使用し、出力の全結合層にはシグモイド関数を使用、最適化手法には Adam、評価関数 MAE (Mean Absolute Error) を用い、学習回数は 100 回とした。

2.2 入力動画の撮影方法

入力を行う際の動画の撮影方法として 2 つの手法を検討する。手法 1 では対象の料理を俯角 0 度で 360 度撮影し、その後、俯角をおおよそ 30 度ずつ同じように撮影しながら上げ、俯角 90 度まで撮影する。(図 1)

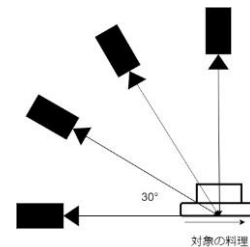


図 1 手法 1 のイメージ

手法 2 では、として対象の料理に対し正面から俯角 0 度で左右 90 度撮影を行い、その後一つ目の手法と同じように俯角をおおよそ 30 度ずつ同じように撮影しながら上げ、俯角 90 度まで撮影する。(図 2)

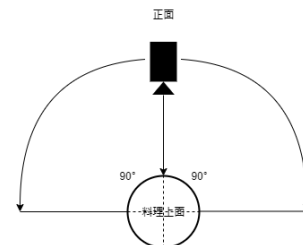


図 2 手法 2 のイメージ

2.3 画像の抽出

2.2 で撮影した料理の動画から魅力度の高い画像の抽出を行うためにまず、魅力度推定器に入力する。を行うための画像を動画から生成する。

画像の生成には OpenCV[3]を使い動画をフレームごとに切り出し、JPEG 形式で保存する。保存した JPEG 形式の画像を魅力度推定器に入力をし、各画像の魅力度を推定し、その中で最も魅力度が高い画像と最も値が低い画像を比較出力した。

3. 評価実験

まず、VGG16 を使った魅力度推定器と精度向上の検証用に、VGG19 を使った魅力度推定器を作成する。次にデータセットに含まれている料理とデータセットに含まれていない料理を撮影、手法 1 と手法 2 に基づいて、動画を生成した。

最後にそれぞれの動画を魅力度推定器に入力し評価を行った。

実験 1

手法 1 で撮影した動画を入力し実験 1 では VGG16 を使った魅力度推定器により評価を行った。図 3 は評価結果を示し、A 画像が魅力度最大、B の画像が魅力度最小を表している。

Photography Assistance System Using An Attractiveness Estimator of Food Photos
Seiya Ohki, Kenneth J Mackin, Yasuo Nagai
(Tokyo University of Information Sciences)



図 3 実験 1 の出力

実験 2

図 4 は VGG16 を使い手法 2 で撮影した動画を入力した結果を示している。

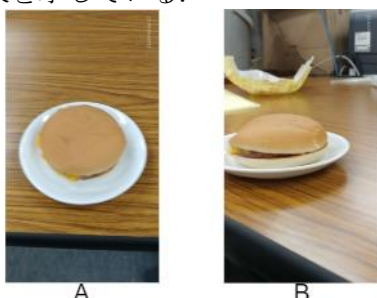


図 4 実験 2 の出力

実験 3

図 5 は実験 3 では VGG19 を使い手法 1 で撮影した動画を入力した結果を示している。



図 5 実験 3 の出力

実験 4

図 6 は VGG19 を使い手法 2 で撮影した動画を入力した結果を示している。

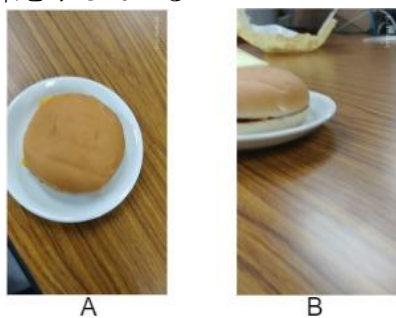


図 6 実験 4 の出力

それぞれの画像の魅力度は表 1 のようになった。

表 1 出力された魅力度

実験番号	魅力度最大	魅力度最大最小
実験 1	0.881235	0.01529771
実験 2	0.95994955	0.02706224
実験 3	0.9891646	0.00118142
実験 4	0.98381305	0.14072469

4. 結論

最適な構図を決めて上手く写真を撮るのは難しいという問題に対して、本研究ではディープラーニングによる料理写真の魅力度推定器を使い、動画から最適な構図の写真を出力する写真撮影補助システムの検討を行った。

その結果、ネットワークの精度に関しては VGG16 と VGG19 どちらも 100 回の学習で高い精度を出すことができた。

また、提案する写真撮影補助システムを利用することで、データセットにある料理の方がデータセットにない料理より良い構図の写真を出力できることがわかった。しかしながら、以下に示すような問題が発生することも明らかになった。

- データセットにない料理では、魅力度が最小値の方が、最大値で出力された写真より良い状態の構図が出力される
- 撮影方法や撮影データのサイズによって結果が変化する

これらの問題に対して、撮影方法を改良することで見切れている写真などを減らす、フレームを数フレーム飛ばして動画の画像化する際に手振れの画像を減らす、本研究で使った以外のネットワークの利用による魅力度推定器の精度の向上などを行うことで対応できると考えている。また、今後は、画像を出力しないでリアルタイムに WEB カメラなどで入力を行い、ユーザーに魅力度を視覚化することで、ユーザーに納得してもらえ写真撮影可能なシステムの実現を検討している。

参考文献

- [1] 佐藤 陽昇, 道満 恵介, 平山 高嗣, 井手 一郎, 川西 康友, 出口 大輔, 村瀬 洋: 畳み込みニューラルネットワークを用いた料理写真の魅力度推定, 信学技報, vol. 117, no. 252, MVE2017-32, pp. 107-111, 2017 年.
- [2] Kazuma Takahashi, Keisuke Doman, Yasutomo Kawanishi, Takatsugu Hirayama, Ichiro Ide, Daisuke Deguchi, Hiroshi Murase: "Estimation of the Attractiveness of Food Photography Focusing on Main Ingredients", Proc. of 9th Workshop on Multimedia for Cooking and Eating Activities (CEA), pp. 1-6, Aug. 2017.
- [3] OpenCV: <https://opencv.org/>