

グリッド変形によるコンテンツ適応型画像リサイズ

Grid-Based Content-Aware Image Resizing

三柴 数*
Kazu Mishiba

池原 雅章*
Masaaki Ikehara

1 まえがき

コンテンツ適応型画像リサイズ手法は、各画素に対して画像の内容に応じた不均一な処理を施すことで視覚的に自然なリサイズを行う技術であり、近年盛んに研究が行われている [1]。提案されている手法の多くは、画像中の重要領域を変化させず、背景などの不要領域に対して画素の削除や縮小、または挿入や拡大を行うことで所望サイズの画像を得る。しかし、所望サイズが重要領域のサイズよりも小さい場合には重要領域の変化が避けられず、リサイズ画像にひずみを生じる原因となる。また所望サイズが画像サイズより大きい場合は不要領域が増加するため、画像全体に占める重要領域の割合が減少する。これは、視覚的に注目を集めるべき領域が目立たなくなることを意味するため、好ましくない。

これらの問題を解決するために、本研究では、重要領域のサイズではなく、そのアスペクト比を保つことによりリサイズを行う手法を提案する。変形の過程で重要領域のアスペクト比を保つことで、様々な表示サイズに対して見た目の違和感を抑えたりリサイズを可能にする。

2 提案法

提案法は、画像をグリッドに投影し、グリッドの変形によりリサイズを行う。変形の過程で重要領域を含むグリッドのアスペクト比を保つことで、見た目に自然なリサイズを行う。提案法の工程は大きく二つに分けられる。一つ目は画像中の重要領域を抽出する工程、二つ目は抽出された重要領域のアスペクト比を保つようにグリッドを変形させる工程である。図1に、提案法を用いたリサイズ工程の概略を示す。

まず一つ目の工程である、重要領域の抽出について述べる。重要領域を抽出するためには、領域の重要度合いを求める必要がある。重要度は人の主観やアプリケーションによって異なるが、本提案法では視覚的に目立つ領域は重要度が高い、と定義する。本提案法における重要領域の抽出では、Achantaら [2] が提案した視覚的に目立つ領域を抽出する手法（以下FSRD法と表記）を用いる。FSRD法は、まず画像中の各領域における視覚的に目立つ度合いを表す視覚的顕著性マップを計算し、これとイメージセグメンテーションを組み合わせることで、視覚的に目立つ領域を抽出する。FSRD法を用いて抽出される領域は重要度があまり高くないまたはサイズが小さい場合があるため、領域の重要度の総和が抽出

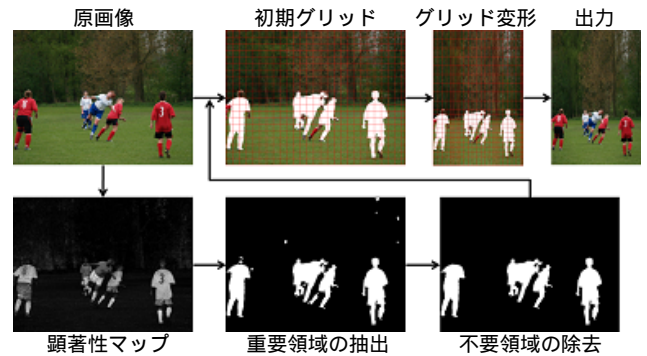


図1: 提案法におけるリサイズのフロー図。

された全領域の重要度の総和の5%以下である領域は、重要領域から除外する。

次に二つ目の工程である、グリッドの変形方法について述べる。ここでは、サイズ $m \times n$ の原画像を $m' \times n'$ にリサイズする場合について考える。提案法は図1に示すように、グリッドを構成する垂直方向のグリッドラインを水平方向に、水平方向のグリッドラインを垂直方向に移動させることでグリッドを変形する。ここで、隣り合う垂直方向のグリッドライン、または水平方向のグリッドラインに挟まれた区間 t に対して、変形前の t の長さを l_t 、変形後を l'_t とする。また、隣り合う垂直方向のグリッドラインに挟まれた区間の集合を T_h 、水平方向を T_v とする。提案法の目標は、 $\sum_{t \in T_v} l'_t = m'$ 、 $\sum_{t \in T_h} l'_t = n'$ 、 $l'_t > 0$ を満たし、最も重要領域のアスペクト比が保たれるような l'_t を求めることである。重要領域 s のアスペクト比が保たれている状態とは、 s を含む区間の変形率 l'_t/l_t が全て等しい場合である。そこでまず、 s を含む各区間の変形率とそれらの平均変形率の差を用いて、 s のアスペクト比が保たれていない度合い $E_A(s)$ を以下の式で定義する。

$$E_A(s) = \sum_{t \in T_s} l_t \left(r_s - \frac{l'_t}{l_t} \right)^2. \quad (1)$$

ここで T_s は、 s を含む水平および垂直方向の区間の集合を、 r_s は T_s に含まれる全ての区間の平均変形率を表す。 T_s の要素数を $N(s)$ としたとき、 r_s は以下の式で表される。

$$r_s = \frac{1}{N(s)} \sum_{t \in T_s} \frac{l'_t}{l_t}. \quad (2)$$

*慶應義塾大学

ここで、画像から抽出された重要領域 s の集合を S とすると、各重要領域における $E_A(s)$ の総和 $E_A = \sum_{s \in S} E_A(s)$ が最小となる l'_t を求めることで、重要領域のアスペクト比を保ったグリッドの変形が可能になる。しかしこの条件では、重要領域を含まない区間に対する制約が存在しないため、その区間の長さが定まらない。そこで、全区間の変形率を評価する以下の関数

$$E_P = \sum_{t \in \{T_v, T_h\}} l_t \left(1 - \frac{l'_t}{l_t}\right)^2 \quad (3)$$

を制約として加え、 $E = E_A + \lambda_P E_P$ を最小にする l'_t を求めることとする。ここで λ_P は E_P に対する重みである。また、重要領域を含んでいない区間であっても、過度の縮小や拡大が行われれば視覚的な違和感を生じる。そこで、変形後の区間長 l'_t が取りうる値を一定の範囲に制限することで、これを防止する。区間長 l'_t が取りうる下限を d_t 、上限を u_t としたとき、これらを画像全体の垂直方向の変形率 $s_v = m'/m$ 、水平方向の変形率 $s_h = n'/n$ を用いて以下のように定義する。

$$d_t = h_d l_t \min\{s_v, s_h\}, u_t = h_u l_t \max\{s_v, s_h\}. \quad (4)$$

ここで h_d, h_u はそれぞれ下限、上限を調整する閾値である。以上を用いて、以下の式で表される最適な変形後の区間長 l'_t を求める。

$$\begin{aligned} & \arg \min_{l'_t} E \\ & \text{s.t. } \sum_{t \in T_v} l'_t = m', \sum_{t \in T_h} l'_t = n', d_t \leq l'_t \leq u_t. \end{aligned} \quad (5)$$

そしてグリッドの変形に対応するように画像の各区間をバイリニア補間を用いて変形することで、リサイズ画像を得る。

3 結果および考察

均一な変形を行うスケーリング手法と提案法を比較し、提案法の有効性を示す。提案法で用いたパラメータは $\lambda_P = 0.1, h_d = 0.4, h_u = 1.5$ であり、グリッドは30ピクセル間隔で区切ったものを用いた。

図2に示すように、提案法は拡大、縮小いずれの場合においても、視覚的に注目を集めると考えられる物体のアスペクト比をなるべく保ったりサイズ画像を生成している。また、拡大時には注目物体の大きさも大きくなるよう変形されており、画像全体に占める重要領域の割合が拡大時に減少してしまうという、従来のコンテンツ適型リサイズ手法の持つ問題点を解決できている。図3上段の結果を見ると、所望サイズが重要領域である車のサイズよりも小さい場合でも、そのアスペクト比をなるべく保ったままリサイズすることで、ひずみの発生を抑えられていることが分かる。図3下段の結果においても、提案法は人物のアスペクト比を保持しているため、視覚的により自然なリサイズ画像を生成できている。しかし、地面に見られる斜めの直線がややひずんでしまっ

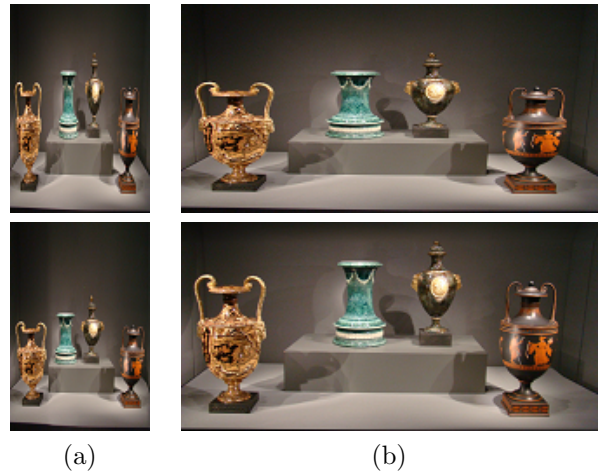


図2: サイズ 600×800 の原画像をスケーリング(上段)と提案法(下段)を用いてリサイズした結果の比較。(a) 原画像の横幅を $1/2$ 倍に縮小。(b) 横幅を 1.5 倍に拡大。



図3: サイズ 512×768 の原画像 (a) をスケーリング (b) と提案法 (c) を用いて横幅を半分にした結果の比較。

ている。今後の課題は、より自然なリサイズ画像を得るために、このような斜めの直線に生じるひずみを抑えることである。

本論文では、Flickr(<http://www.flickr.com/>) においてクリエイティブ・コモンズ・ライセンスに基づいて公開されている画像をシミュレーションに用いました。画像を公開して下さった以下の方々に感謝致します。dmondark(Decorative vases), gwenael.piaser (two little tourists), Ingy The Wingy (soccer), twicepix(car). (敬称略)

参考文献

- [1] D. Vaquero, M. Turk, K. Pulli, M. Tico, and N. Gelfand, "A survey of image retargeting techniques," Proc. SPIE Applications of Digital Image Processing XXXIII, 2010.
- [2] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," Proc. IEEE Conf. CVPR, 2009.