

論文 / 著書情報
Article / Book Information

題目(和文)	任意形の表示領域のためのイメージリターゲッティングに関する研究
Title(English)	
著者(和文)	野中敬介
Author(English)	Keisuke Nonaka
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9549号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:羽鳥 好律,長橋 宏,小林 隆夫,小池 康晴,佐藤 いまり,宮田 高道
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第9549号, Conferred date:2014/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

任意形の表示領域のための
イメージリターゲティング
に関する研究

[学 位 論 文]

2014年3月

指導教員 羽鳥 好律 教授

提出者 大学院 総合理工学研究科
物理情報システム専攻
野中 敬介

目次

第1章	序論	1
1.1	研究の背景	1
1.2	イメージリターゲッティングにおける問題の所在	4
1.3	研究の位置づけ	6
1.4	論文の構成	9
第2章	既存研究	10
2.1	まえがき	10
2.2	イメージリターゲッティングにおける ROI 決定法の位置づけ	11
2.3	ピクセルベースアプローチによる既存のイメージリターゲッ ティング	12
2.3.1	Seam Carving	12
2.3.2	Image Summarizing via Bidirectional Similarity	14
2.4	ワーピングベースアプローチによるイメージリターゲッティング	17
2.5	むすび	19
第3章	提案手法1:ピクセルベースの任意形イメージリターゲッティ ング	21
3.1	まえがき	21
3.2	提案手法1	22
3.2.1	画像中の重要領域の抽出	23
3.2.2	表示領域に応じた重要領域の移動	24
3.2.3	初期推定画像を用いた任意形画像の生成	25
3.3	実験および考察	26
3.3.1	Bidirectional Similarity と主観評価との相関	26
3.3.2	比較実験	27
3.3.3	パラメータ設定における頑健性に関する考察	28

3.3.4	考察	29
3.4	むすび	32
第4章	提案手法2：ワーピングベースの任意形イメージリターゲッ ティング	35
4.1	まえがき	35
4.2	準備	35
4.3	提案手法2	36
4.3.1	概要	36
4.3.2	ROI の抽出	39
4.3.3	格子点に対するエネルギー汎関数	39
4.4	アルゴリズム	44
4.4.1	Alternative Directions Method for Multipliers (ADMM)	44
4.4.2	ADMM の拡張	45
4.4.3	提案手法2における ADMM	47
4.5	実験	48
4.5.1	実験環境	48
4.5.2	考察	50
4.6	むすび	55
第5章	結論	57
	謝辞	60
	参考文献	62
	発表論文	66
	付録I	68
I.1	定理4.4.3 の証明	68
I.2	定理4.4.4 の証明	69

目次

1.1 一般的な画像処理を用いた画像表示を行った際に想定される問題例	2
1.2 従来のイメージリターゲッティングによる出力の一例（文献[1], Figure 8(a), (f)より引用）	2
1.3 本研究の対象とする課題	4
1.4 インペインティングとリターゲッティングの対象とする問題の違い	5
2.1 Seam Carving の一例（文献[1], Figure 1より引用）	13
2.2 Completeness と Coherence	16
2.3 ブロックマッチングを用いた画素値の更新	16
2.4 Simakov らの手法[2]の概要	18
2.5 Simakov らの手法の一例（文献[2], Figure 5より引用）	18
2.6 Feature-aware texturing の一例（文献[3], p. 230, Figure 4より引用）	19
3.1 提案手法1の概要	22
3.2 (a) 原画像, (b) 原画像を任意形の表示領域で表示した画像, (c) ~ (e) 既存手法を用いて, 任意形の表示領域に外接するように変換した画像, (f) 提案手法1結果画像, (左から右に [bison, fish, dolphin, human-bird], 黒色: 非表示領域)	33
3.3 (a) 原画像, (b) 重要領域抽出結果, (c) ~ (f) 重要領域移動結果, 黒色: 値が未定の画素, 赤色: 非表示領域, 各画像（黒色, 赤色領域を含む）はいずれも原画像と同サイズ	34
4.1 提案手法2の概要	38
4.2 ROI 抽出の一例 (a) 原画像, (b) 顕著性マップ, (c) ROI 候補画素（白色）, (d) ROI	39

4.3	A_i の概略	42
4.4	$W_{i,j}$ の概略	42
4.5	ハート形の表示領域における既存手法との比較（画像の名称は上から下へ順に, “Fish (heart)”, “Cars (heart)”, “Deer (heart)”, “Child (heart)”であり, 各行の左上の画像が原画像である.) . . .	51
4.6	多様な形状の表示領域における既存手法との比較（画像の名称は上から下へ順に, “Fish (diamond)”, “Cars (spade)”, “Dolphins (club)”であり, 各行の左上の画像が原画像である.)	52
4.7	従来までの対象である矩形の表示領域における既存手法との比較（画像の名称は上から下へ順に, “Deer (rect.)”, “Rope skipping (rect.)”, “Cliff (rect.)”であり, 各行の左上の画像が原画像である.)	53
4.8	閾値 κ による結果の違い（図 4.5 の結果は (c)）	55
5.1	任意形イメージリターゲッティングのアプリケーション応用例 .	58

表 目 次

1.1	提案手法の位置づけ	8
3.1	評価実験結果 1 (bison [250 × 188 pixel])	29
3.2	評価実験結果 2 (fish [320 × 240 pixel])	30
3.3	評価実験結果 3 (dolphin [320 × 240 pixel])	30
3.4	評価実験結果 4 (human-bird [300 × 225 pixel])	30
3.5	異なるパラメータ (α, β, γ) における Bidirectional Similarity の損失量 d . ($\gamma = 1.0 - (\alpha + \beta)$)	31
4.1	各画像における閾値 κ	50

第1章

序論

1.1 研究の背景

近年，携帯電話やノートパソコン，液晶テレビをはじめ，デジタルフォトフレームやデジタルサイネージなど，デジタル画像を扱う多様なデバイスが生活の中で幅広く利用されている．これらの既存のデバイスのほとんどは，それぞれが異なる目的を有して設計されているにも関わらず，筐体が長方体であるという共通の特徴を持っており，そのためにデジタル画像を提示するディスプレイ部（以下，表示領域）も同様に，長方形（以下，矩形）であることが多い．これら多様なデバイスにおける矩形の表示領域上で，単一の矩形画像を表示した場合には，様々な問題が起こる．

例えば，入力画像とアスペクト比の異なる表示領域上において，bicubicなどの一般に広く用いられている拡大縮小処理による画像表示を行った場合，画像と表示領域のアスペクト比の相違から，出力される画像のオブジェクトが大きく歪むことが予想される（図 1.1(a)）．また，サムネイルなどの小さなサイズの表示領域上に対して，同様の縮小処理を用いて画像を表示した場合，視覚的に重要な要素の細部が損なわれ画像の内容をユーザーが理解しづらくなる，といった問題が起きる（図 1.1(b)）．

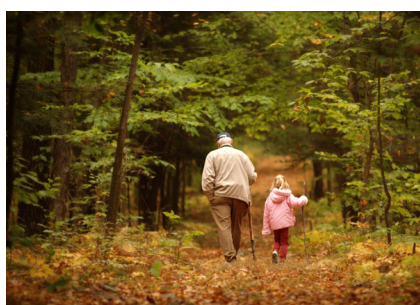
このような画像と表示領域の差異に起因する問題を解決する方法として，表示領域に応じた適応的な表示を可能とするイメージリターゲットイングと呼ばれる手法が数多く提案されている．ここで，適応的な表示とは，入力画像を異なるサイズ，アスペクト比，解像度の画像出力領域（以下，表示領域）で表示する際に，そのコンテンツ（ライン，顔，オブジェクトの形



(a-1) 原画像



(a-2) 出力画像



(b-1) 原画像



(b-2) 出力画像

図 1.1 一般的な画像処理を用いた画像表示を行った際に想定される問題例

状など)の歪みや極端な拡大縮小を抑えながら、違和感のない画像表示を行うことを意味する(図 1.2)。実際に、矩形の表示領域を対象とした従来のイメージリターゲッティング(以下、矩形イメージリターゲッティング)[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]では、「原画像内の視覚的に顕著な領域(Region of Interest, ROI)の形状を保存しながら表示領域に収まるように移動」させつつ、同時に「画像全体の歪みを抑え違和感のない画像生成を行う」といった問題を解いていると捉えることができる。



(a) 原画像



(b) 出力画像

図 1.2 従来のイメージリターゲッティングによる出力の一例(文献[1], Figure 8(a), (f)より引用)

一方で、デジタルフォトフレームやデジタルサイネージといったデバイスについては、その使用目的がインテリアや広告であることから、今後はよりデザイン性、宣伝性を重視した矩形に限らない（以下、任意形）筐体をもつ製品の需要が高まることが予想される。また、近年では Google Glass[15] に代表されるようなウェアラブルデバイスも注目されており、それらについても同様の理由から、任意形の筐体をもつものが適していると考えられる。このように、これまでの「ディスプレイデバイスは矩形である」という形状の制約を取り除くことにより、新しい様々な応用が生まれることが期待できる。このような背景から、文献[16, 17, 18, 19]では任意形のディスプレイデバイスの製造技術が提案されている。これらの任意形の製品の表示領域としては、従来の矩形の表示領域ではなく、筐体と同様の任意形の表示領域を採用することがより有用であるといえる。

しかしながら、任意形の表示領域上での画像表示に対して、単純に従来の画像処理技術を適用した場合、様々な問題が生じる。例えば、任意形の表示領域全体を使用するように矩形画像を表示した場合、表示領域と画像の形状の不一致のために、表示できない矩形画像の領域に含まれていた重要な要素が損なわれる可能性がある（図 1.3(a)）。一方で、単純な画像縮小処理によって矩形画像全体を任意形の表示領域上で表示する場合には、表示領域の一部を使用することができないだけでなく、本来の意図であるデザイン性をも損なう結果となる（図 1.3(b)）。また、従来の矩形の表示領域を対象としたイメージリターゲッティング（以下、矩形イメージリターゲッティング）は、そのアルゴリズムの設計において、「出力表示領域の形状が矩形である」という制約を大きな手がかりとして用いているため、単純に任意形の表示領域を対象とする手法として拡張することはできない。以上の考察より、従来の矩形画像を任意形の表示領域上で適応的に表示するためには、何らかの新しい画像処理手法が必要となる。

そこで、本研究では任意形の表示領域のためのイメージリターゲッティング（以下、任意形イメージリターゲッティング）を提案する。提案する任意形イメージリターゲッティングでは、従来の矩形イメージリターゲッティングと同様に、「ROI の形状を保存しつつ可能な限り表示領域内に収める」とことと、「画像全体の歪みを抑え違和感のない画像生成を行う」とことの両立を目指す。

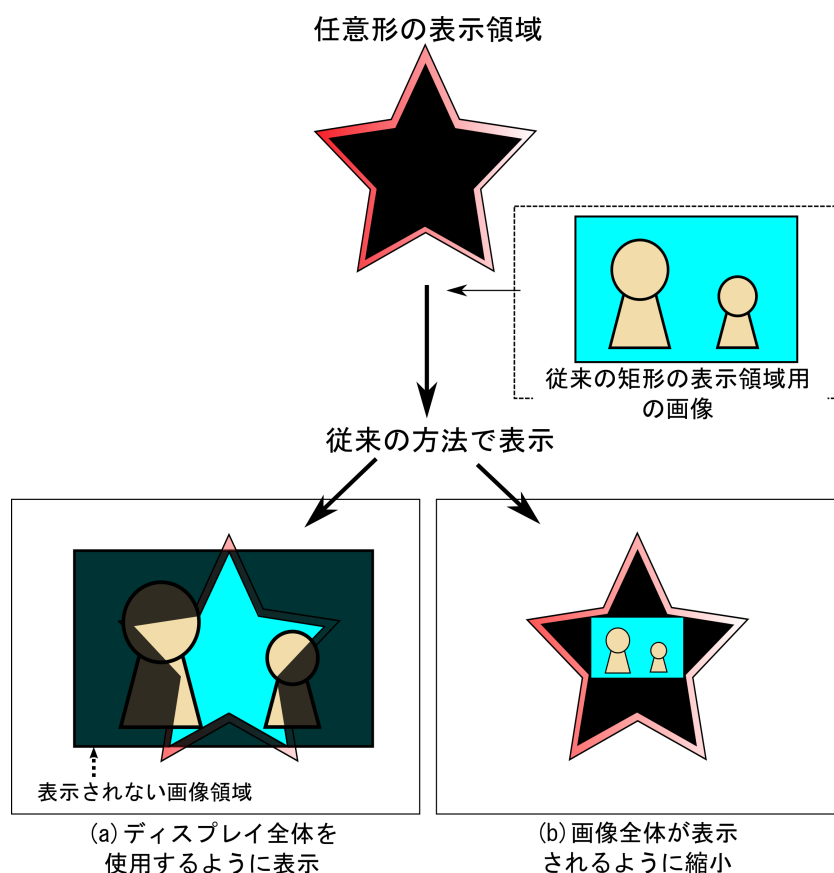


図 1.3 本研究の対象とする課題

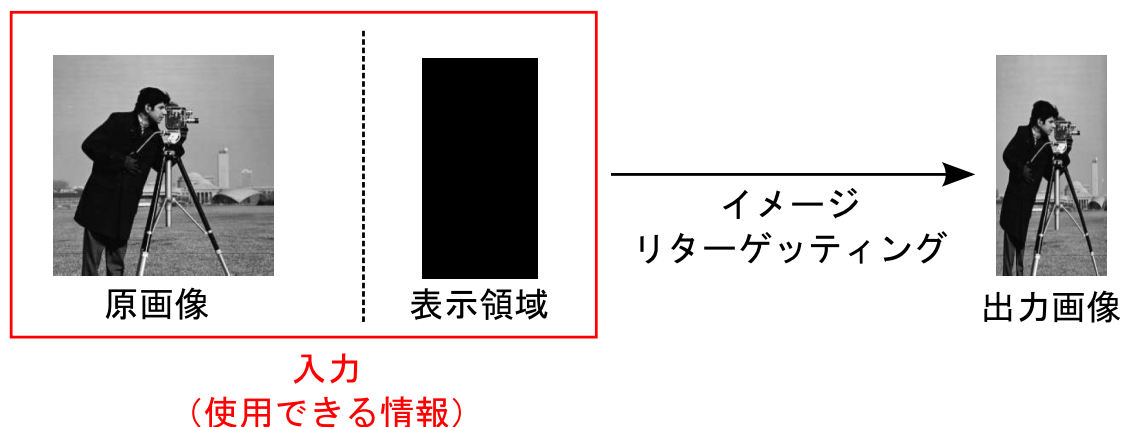
1.2 イメージリターゲッティングにおける問題の所在

イメージリターゲッティングは、原画像のもつ画素値や構造を、画素値が未定である出力表示領域にどのようにマッピングするか、を解く問題であるといえる。以下、任意形の表示領域に表示される画像を任意形画像と定義し、矩形画像の画素値を任意形画像の画素値へマッピングする操作を画像変換（または単に変換）と呼ぶ。

イメージリターゲッティングと同様に画素値が未定の領域の画素値を決定する研究として、画像の欠損領域（画素値未定）を修復するイメージインペインティングが挙げられる（図 1.4(a)）。このインペインティングでは、入力である原画像が欠損しているため、原画像のもつすべての画像情報を用いることはできない。しかしながら、欠損領域の周辺の画素値や、オブジェクト境界の連結性などの画像の構造を手がかりとして、望ましい画像を得ること



(a) イメージインペインティング (文献 [20], p. 10, Figure 6 より引用)



(b) イメージリターゲッティング (文献 [3] の手法による処理結果)

図 1.4 インペインティングとリターゲッティングの対象とする問題の違い

が可能である．これに対して，イメージリターゲッティングにおいて入力として使用できる情報は，原画像そのものと出力表示領域の画素の座標（表示領域の形状）のみであり，インペインティングの例に比べて出力側に関する情報が非常に少ない（図 1.4(b)）．このため，どのように画像を変換すべきかという方針を立てることがより一層難しくなる．

この問題に対して，従来の矩形イメージリターゲッティングでは，原画像と出力表示領域がともに矩形であるという形状の制約を用いることで，ROIの形状を保存しつつ原画像のもつ内容を損なわないような画像出力を可能としている．つまり，「画像境界に近い画素は，同様に出力画像においても対応する画像境界に近い」というような，境界からの画素の位置関係による制約が，変換の大きな手がかりとなっているといえる．特に，ROIに含まれる画素につ

いては、可能な限り形状を保存するという目的のため、「ROI内の画素の隣接関係を崩さない」といった制約も併せて使用することができ、そこから周辺の画素値を推定することが可能である。これは、インペインティングにおいて欠損領域の周辺の情報を使用している例と類似している。以上の議論から、矩形イメージリターゲッティングは、「表示領域の形状の制約を用いて原画像全体の構造を考慮しつつ、ROIに含まれる画素の座標を優先的に決定する問題を解く」、という枠組みに帰着できることがわかる。

一方で、本研究の対象である任意形の表示領域においては、矩形の原画像と表示領域の形状の対応関係は自明ではない。そのため、従来のイメージリターゲッティングが手がかりとしてきた画像境界の対応関係を使用することはできず、対象とする問題はより困難なものとなる。しかしながら、任意形イメージリターゲッティングにおいても、原画像のROIの位置関係やそれに含まれる画素の隣接関係は使用可能である。そのため、矩形イメージリターゲッティングと同様に、「ROIに含まれる画素の、優先的な表示領域上の座標決定」と、「原画像のもつ構造を保った画像出力」とを同時に行うといった問題の本質は変わらない。

1.3 研究の位置づけ

従来のイメージリターゲッティングにおける最も代表的な手法のひとつとして、2007年にAvidanらによって提案されたSeam Carving[1]が挙げられる。Seam Carvingは、seamと呼ばれる画素の集合を定義した上で、原画像の構造を最も損なわないseamの削除や複製を繰り返すことによって、図1.2のような全体として違和感のない（原画像の構造を損なわない）画像を生成することを可能としている。この手法の最大の特徴は、seamの定義および画像の生成法が非常に簡易なものであり、拡張性や応用性に優れる点である。

Seam Carvingの後に提案されたイメージリターゲッティング手法としてSimakovらのImage Summarizing via Bidirectional Similarity[2]が挙げられる。Seam Carvingが処理全体を通してどのような評価基準を最小化しているかが不明瞭であるのに対して、Simakovらの手法では原画像と出力画像との画素値を用いて算出される客観評価指標Bidirectional Similarityを導入し、それを最小化することを目的としており、その点がこの手法の最大の特徴であるといえる。

これら2つの手法の共通点は、画素ごとに定義される何らかのエネルギー汎関数の最小化問題を解くことで、明示的に未定の画素値を決定する手法である点である。以降では、このような手法をピクセルベースアプローチと呼ぶ。同様に、画素値ごとのエネルギー汎関数の最小化を試みている手法としては、PritchらのShift-map[6]などが挙げられる。

ピクセルベースアプローチの利点としては、Seam Carvingのように簡易なアルゴリズムから最終的な出力画像を得ることができる点や、処理過程において出力画像に対する直感的な理解が可能であるといった点から、広い拡張性、応用性をもつことである。また、Image Summarizing via Bidirectional Similarityのように、出力画像に対して客観的な評価基準を定義し、その評価基準における最小化が可能であるという点である。

しかしながら、上記のピクセルベースアプローチのイメージリターゲッティングは、定義されるエネルギー汎関数が非凸なものとなる（関数の凸性については第4章において定義する）。これは、多くの場合、最適化問題の中で扱う変数の数が画素数以上の膨大な数になり、凸な最適化問題として定式化することが困難となるためである。一般に、関数が非凸な問題の大域最適解を導出することは非常に困難な問題である。そのため、すべてのピクセルベースのアプローチは、発見的なアルゴリズムによりエネルギー汎関数の最小化を試みているものの、最終的な出力画像に対する最適性は保証されず、局所的な最小解を得ることに留まっている。また、画素ごとに出力する値を決定するため、画像の歪みや断裂を生じやすいことも、ピクセルベースの欠点である。

これに対して、原画像を格子状に分割した際に得られる格子点座標に対してエネルギー汎関数を定義し、その最小解から画像を得る、ワーピングベースと呼ばれるアプローチに分類される手法[3, 7, 8, 9, 10, 11, 12, 13, 14]がある。このワーピングベースアプローチは、ピクセルベースの手法に比べて出力画像に不自然な歪みや断裂が生じにくいことが特徴であり、近年盛んに研究対象とされている。また、このアプローチでは、最適化問題で扱う変数が格子点の座標に相当するため、ピクセルベースの手法と比べて一般に変数の数が大幅に少ない。そのため、格子点の座標決定問題は凸なエネルギー汎関数の最適化問題として定式化可能であり、そのエネルギー汎関数において大域最適性をもつ画像を得ることが可能である点もワーピングベースアプローチの特徴である。このように、ピクセルベースアプローチの手法の欠点を補うワーピングベースアプローチの手法であるが、処理の過程において最終的にどの

表 1.1 提案手法の位置づけ

出力画像の最適性 (エネルギー汎関数の凸性) 表示領域の形状	局所最適 (非凸)	大域最適 (凸)
矩形	ピクセルベース の既存手法 [1, 2, 4, 5, 6]	ワーピングベース の既存手法 [3, 7, 8, 9, 10, 11, 12, 13, 14]
任意形	第3章	第4章

ような画像が生成されるか理解することが難しく、拡張性、応用性に乏しいといった面もある。

以上をまとめると、従来のイメージリターゲッティングは、出力表示領域における未定の画素値を明示的に決定する「ピクセルベース」と呼ばれるアプローチと、原画像を格子状に分割した際に得られる格子点の座標を決定する問題を解く「ワーピングベース」と呼ばれるアプローチとに分類することができる。また、ピクセルベースのイメージリターゲッティングは、手法ごとに定められたエネルギー汎関数における局所最適解として画像を出力するものであり、ワーピングベースのイメージリターゲッティングは、エネルギー汎関数における大域最適解として画像を出力するものである。また、それぞれのアプローチにおいて、上述のように様々な利点、欠点が存在する。

このように、2つの分類が可能であるイメージリターゲッティングであるが、すべての手法において何らかの評価基準におけるエネルギー汎関数の最小化を行うという枠組みは共通している。そのため、任意形の表示領域を対象とした際に、ユーザーの望む要素をどのようにエネルギー汎関数として設計し、解を導出するかという点が本研究の焦点となる。本論文では、表 1.1 のように、既存のイメージリターゲッティングの2つの分類に対応する2つの手法を提案することで、任意形の表示領域におけるエネルギー汎関数の設計方法および解の導出方法（画像の生成アルゴリズム）を提案し、任意形イメージリターゲッティングに対する包括的な解決を試みる。

第3章にて提案する手法は、既存のピクセルベースの矩形イメージリター

ゲッティング手法において提案された客観評価指標に基づく、任意形イメージリターゲッティングである。この手法は、既存のピクセルベースの手法と同様に、非凸なエネルギー汎関数を局所最小化した画像を出力する。

これに対して、第4章にて提案する手法は、既存のワーピングベースの枠組みに則り、凸なエネルギー汎関数を定義し、それに対して凸最適化の理論を用いることで、大域最適解となる画像を出力する任意形イメージリターゲッティングである。

1.4 論文の構成

本論文は以下のように構成されている。まず、第2章では従来の矩形イメージリターゲッティングを、そのアプローチによって分類し、それらの詳細と問題点について述べる。次に第3章において、ピクセルベースの任意形イメージリターゲッティング手法を提案する。続く第4章において、ワーピングベースの任意形イメージリターゲッティング手法を提案する。最後に、第5章において、本研究の対象とする任意形リターゲッティングのアプリケーション応用例を示すとともに、提案手法の今後の課題について議論を行うことにより、本論文の統括とする。

第2章

既存研究

2.1 まえがき

前章で述べたように、既存のイメージリターゲットングはそのアプローチに基づき、ピクセルベースアプローチ [1, 2, 4, 5, 6] と、ワーピングベースアプローチ [3, 7, 8, 9, 10, 11, 12, 13, 14] とに大別できる。

ピクセルベースアプローチは、原画像の画素値を用いて出力表示領域の画素値を直接決定する手法であり、比較的シンプルなアルゴリズムで構成されているため、実装や応用が容易であるという特徴がある。しかしながら、画素（もしくはブロックと呼ばれる矩形の小領域）単位で画像を扱うため、それらの境界において重要なオブジェクトの不自然な歪みや断裂、ぼけを生じやすい。また、すべての画素の値を直接決定するため、膨大な数の変数を計算対象とする場合が多く、それらを決定するためには非凸なエネルギー汎関数の最適化問題を解く必要がある。このことは、最終的に出力される値に対して、何らかの評価基準における大域最適性が保証されないことを意味している。

一方、ワーピングベースアプローチは、入力画像上に設定した均一な格子点の表示領域上での移動先を決定し、その座標を用いて画像を出力する手法である。格子点の座標は明示的に出力画像の特徴を決定するものではないため、処理の過程およびアルゴリズムの構成において、直感的な理解や応用が困難であるものの、出力される画像には、オブジェクトの断裂が少なく、ぼけも生じにくいという特徴がある。一般に、格子点の総数は表示領域上の画素数より大幅に少なく、格子点の座標決定問題は、凸な最適化問題として定義することが可能である。実際に、既存のワーピングベースアプローチは、格

子点の座標決定において凸な最適化問題を解いている。また、ピクセルベースの手法が、それぞれの手法ごとにまったく異なる独自のアルゴリズムを用いていることに対して、ワーピングベースの手法は格子点の座標を決定するという枠組みは共通しており、手法による差異は座標の決定方法のみである。

ここで、すべての既存手法は、ピクセルベース、ワーピングベースといったアプローチの分類によらず、出力画像が矩形であるという形状の制約を用いており、単純に任意形の表示領域に適用することはできない。

以降の節では、まず多くのイメージリターゲッティングにおいて必要とされる ROI の決定法の位置づけについて述べる。その後、従来のピクセルベース、ワーピングベースの代表例について詳細に述べ、任意形の表示領域を対象とした際の問題点を明らかにする。

2.2 イメージリターゲッティングにおける ROI 決定法の位置づけ

ROI (Region of Interest) とは画像中の注目されやすい領域のことであり、イメージリターゲッティングにおいては、原画像内における視覚的に顕著な領域の形状を保存する目的のために、ROI を決定する必要がある。そのため、イメージリターゲッティングにおける ROI の決定問題は、最終的な結果画像の良し悪しに大きく影響を与える重要な要素である。当然ながら原画像内に対して誤った ROI が決定された場合、出力画像は望ましいものにはならない。

この ROI の決定問題に対して、多くのイメージリターゲッティング手法では、ユーザーが画像を見た際の視覚刺激をスカラー量として画像で表現した顕著性マップ [21, 22, 23, 24, 25, 26] が用いられる。これら顕著性マップが、ユーザーが画像を見た際にどの領域に注目しやすいかを自動的に推定できる簡易な手法であるという特徴をもつためである。また、その他のイメージリターゲッティング手法の中には、原画像内のエッジの強度を ROI として採用したり、領域分割やユーザー入力を補助的に利用したりするものもある。

このように、イメージリターゲッティングにおける ROI 決定の目的のためには様々な方法が用いられているが、多くの手法においてその方法は厳密に規定されてはいない。これは、例えばユーザー入力などを用いて画像ごとに望ましい ROI を細かに設定し、結果画像をより良くすることはできるが、実用

性の面から厳密な ROI 設定に対してのみ良い結果を生成するリターゲッティング手法は望まれないためである。つまり、イメージリターゲッティングにおいて ROI の決定問題というのはその結果を良くするために補助的に利用されるものであり、研究対象となるのは「おおよそ適切な ROI が与えられた場合に、どのようなエネルギー汎関数を設計し、どのようなアルゴリズムを用いてよりユーザーが望む画像を生成するか」という点である。

以上の議論から、本論文における研究対象は、ROI が決定された後のリターゲッティング処理アルゴリズムの提案であることを明記しておく。

2.3 ピクセルベースアプローチによる既存のイメージリターゲッティング

本節では、ピクセルベースアプローチのイメージリターゲッティングの代表例として、Seam Carving[1, 4, 5] と Image Summarizing via Bidirectional Similarity[2] について述べる。

2.3.1 Seam Carving

Seam Carving[1] は 2007 年に Avidan らによって提案されたイメージリターゲッティング手法であり、非常に簡易な枠組みによりアルゴリズムが構成されているため、高い拡張性をもつことが特徴である。現在でも、Seam Carving の拡張や応用を行った手法[27, 28, 29, 30]が多く提案されていることから、この手法の有用性がわかる。

この手法は、seam と呼ばれる原画像を横断（または縦断）する 8 連結の画素列（図 2.1 参照）のうち、文献[1, 4, 5]によって定義されたエネルギーが最小のものを探索し、それを反復ごとに削除、複製することで画像のサイズを変更している。いま、画像を高さ n 、幅 m とすると、垂直方向の seam は以下のように定義される。

$$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, \text{ subject to } \forall i, |x(i) - x(i-1)| \leq 1 \quad (2.1)$$

ここで、 $x()$ は写像 $x() : \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ である。また式 (2.1) は、seam 中の任意の画素において、その直上および直下の画素は左右に最大 1 画素しか変



(a) 原画像（赤色の画素は選択された seam）

(b) 出力画像

図 2.1 Seam Carving の一例（文献 [1], Figure 1 より引用）

位できない，ということを意味している．同様に， $y() : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$ としたとき，水平方向の seam は以下のように定義される．

$$s^y = \{s_j^y\}_{j=1}^m = \{(j, y(j))\}_{j=1}^m, \text{ subject to } \forall j, |y(j) - y(j-1)| \leq 1 \quad (2.2)$$

さらに，ある画素 $p \in \mathbb{R}^2$ に対して計算されるエネルギー $e(p)$ を用いて，ある seam, s のエネルギーは以下のように定義される．

$$E(s) = \sum_{i=1}^{n_s} e(s_i) \quad (2.3)$$

ここで， s_i は s に含まれる画素をあらわし， n_s はその総数である．また，式 (2.4) のように画像内においてエネルギーが最小となる seam を s^* とする．

$$s^* = \arg \min_{s \in \mathbb{R}^{2n_s}} E(s) \quad (2.4)$$

この s^* は，削除された際に， $e()$ の表す画素の重要度の評価基準において画像のもつ構造に最も影響を与えない seam であるといえる．

以上のように定義した seam を用いて，画像変換を行う．まず， $n \times m$ のサイズの画像を用いて， $n \times m'$ ($m' < m$) のサイズの画像を生成することを考える ($m - m' = c$ とする)．いま，原画像 S に対して，エネルギーが最小となる垂直方向の seam, s^* を計算しその画素を削除する．この操作により画像幅が 1 [pixel] 縮小された画像を得ることができる．得られた画像に対してエネルギーを再計算し，その画像から再度垂直方向の s^* を選択し，削除するという一連の操作を c 回繰り返す．以上の操作により， c 本の seam を削除することとなり，幅が m' である画像を得ることができる．

次に、入力画像をより大きいサイズに変換する方法について述べる。画像を拡大するには逆の操作，すなわち選択された seam を複製し、画像内に挿入する操作を行えばよい。

しかし、拡大処理は縮小処理とは若干異なる操作を必要とする。例えば、画像 S の幅を 1 画素分拡大する場合は、 S 内において最もエネルギーが小さい垂直方向の seam を選択し、複製することで所望の画像を得ることができる。しかし、得られた画像を用いて同様の操作を行い、さらに 1 画素分拡大しようとした際には、前回の処理と同じエネルギー最小の seam が選択されてしまう。このように、縮小処理と同様に seam の選択と複製を繰り返すだけでは、同一の seam の複製が繰り返し行われるだけであり、結果として不自然な画像が生成される。これを解消するためには、まず画像 S に対して Seam Carving の縮小処理を適用し、 k 画素縮小した画像を生成する。この時に、選択されたすべての seam の座標をあらかじめ保存しておく。そして、 S においてそれらの seam が該当する位置の画素を順次複製することで、 k 画素分だけ幅の拡大された画像を得ることができる。画像の高さを変える場合に関しても、同様である。

ここまで挙げた、Seam Carving を用いた拡大縮小処理はすべて自動で行われる。文献 [1] では、さらにユーザー入力を用いることにより、画像内の任意の領域を削除しつつ自然な画像を生成することが可能であるとしている。まず、ユーザー入力により、画像内において削除したい領域を選択し、その領域に含まれる画素のエネルギーを小さな値にする。また逆に、画像内において削除したくない領域を選択し、その領域に含まれる画素のエネルギーを大きな値にする。この操作を行った後、上記 Seam Carving による縮小拡大処理を用いることにより、ユーザーにとってより望ましい画像を生成することが可能である。

以上のように、Seam Carving による画像変換は、すべて seam の削除と複製の繰り返しにより実現される。しかしながら、この seam の設計は、いずれの位置においても画像幅（または高さ）が変化しないという矩形の表示領域の形状の特徴を活かしたものであるため、Seam Carving を任意形の表示領域に対して適用することは困難である。

2.3.2 Image Summarizing via Bidirectional Similarity

Seam Carving [1] と同様に広く知られているピクセルベースのイメージリターゲッティングとして、Simakov らの Image Summarizing via Bidirectional

Similarity[2] が挙げられる．この手法では変換前，変換後の二つの画像の間における画像内容の類似度 Bidirectional Similarity を定義し，その損失量を最小化するヒューリスティックな手法を提案していることが特徴である．Bidirectional Similarity の損失量は，変換の際の入力画像，出力画像候補において，双方向からのブロックマッチングを行うことで算出される，Completeness, Coherence と呼ばれる二つの指標によって定義される．

いま，入力画像を S (Source)，出力画像候補を T (Target) とそれぞれ呼ぶこととする．また，それらの画像に含まれる，重複を許した固定サイズの矩形小領域をパッチと呼び， S および T に含まれるパッチをそれぞれ P, Q で表す．さらに S, T の取りうるパッチの総数を N_S, N_T とすると，Completeness, Coherence のそれぞれの損失量 d_{comp}, d_{cohere} ，および Bidirectional Similarity の損失量 d は以下の式で定義される．

$$d_{comp} = \frac{1}{N_S} \sum_{P \in S} \min_{Q \in T} D(P, Q) \quad (2.5)$$

$$d_{cohere} = \frac{1}{N_T} \sum_{Q \in T} \min_{P \in S} D(Q, P) \quad (2.6)$$

$$d = \varepsilon \cdot d_{comp} + (1 - \varepsilon) \cdot d_{cohere} \quad (2.7)$$

ここで， $D(\cdot)$ は $L^*a^*b^*$ 色空間上における，二つのパッチの対応する画素値の二乗誤差の平均であり，パッチ間の距離を表す．また，式 (2.7) における ε は Completeness, Coherence の重みを決定するパラメータである．本稿ではこれ以降，文献 [2] で用いられている値 $\varepsilon = 0.5$ を使用する．

次に，これらの二つの指標が具体的にどういった画像内容を表しているのかについて考察する．式 (2.5), (2.6) より，Completeness は「 S のもつ画像特徴を T がどれだけ含んでいるか」を表し，Coherence は「 T が S に存在しない画像特徴を含んでいないかどうか」を表す指標であることがわかる (図 2.2)．このことから，Bidirectional Similarity は，同一のパターンを持つテクスチャ・オブジェクトなどの繰り返し要素に比べ，画像内のユニークなオブジェクトやそれらの構造を，より重要な要素として評価する画像評価指標であるといえる．

上記の画像評価指標を考慮した上で，Simakov らの手法では，以下のように画像変換を行っている．

まず，入力画像 S を既存の bicubic 法等を用いてわずかに (例えば，0.95 倍に) 縮小したものを $T_1^{(0)}$ とする．次に，この $T_1^{(0)}$ に含まれる画素 q の値 $T(q)$ を，ブロックマッチングを用いて以下のように更新する (図 2.3)．

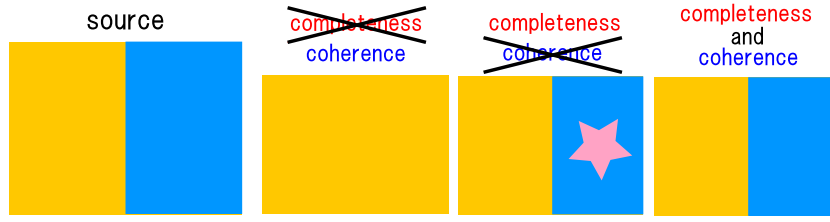


図 2.2 Completeness と Coherence

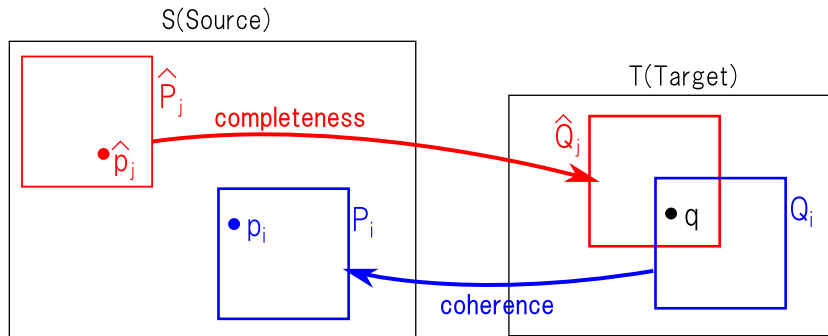


図 2.3 ブロックマッチングを用いた画素値の更新

$$T(q) = \frac{\frac{1}{N_S} \sum_{j=1}^n S(\hat{p}_j) + \frac{1}{N_T} \sum_{i=1}^m S(p_i)}{\frac{n}{N_S} + \frac{m}{N_T}} \quad (2.8)$$

まず, $p_i (i \in \{1 \dots m\})$ について述べる. 画素 q を含むパッチ Q_i に対して, $D(\cdot)$ が最小となる S 内のパッチ P_i を考える. この時 p_i は, パッチ Q_i における画素 q の位置と対応するパッチ P_i 内の位置の画素を表している. また, m は画素 q を含む取りうるパッチの総数である (例えば, 8×8 画素のパッチの場合, $1 \leq m \leq 64$).

次に, $\hat{p}_j (j \in \{1 \dots n\})$ について述べる. いま, S 内のあるパッチ P に対して $D(\cdot)$ が最小となる $T_1^{(0)}$ 内のパッチを探索した際に, 着目している画素 q を内部に含むパッチ Q が選択されたとする. このような T 内のパッチ Q を $\hat{Q}_j$, 対応する S 内のパッチ P を $\hat{P}_j$ と呼ぶ. この時 $\hat{p}_j$ は, パッチ $\hat{Q}_j$ における画素 q の位置と対応するパッチ $\hat{P}_j$ 内の位置の画素を表している. また, n は選択された画素 q を内部に含むパッチ $\hat{Q}_j$ の総数である ($0 \leq n \leq N_S$). ただし, 画素 q によっては選択されるパッチ $\hat{Q}_j$ の総数 n は 0 となりうる. この時, $T(q)$ の更新式

は以下のとおりである．

$$T(q) = \frac{\sum_{i=1}^m S(p_i)}{m} \quad (2.9)$$

また， $S()$ は S 内における画素の値である．

この画素値更新を一度 $T_1^{(0)}$ に含まれる全画素に対して行い，その結果を新たに $T_1^{(1)}$ とする．続いて，この $T_1^{(1)}$ に対して S との間でブロックマッチングを行う．このとき，選択されるパッチ P_i ， $\hat{P}_j$ が更新される．新しいパッチ P_i ， $\hat{P}_j$ について式 (2.8) を用いて画素値更新を行い，これを $T_1^{(2)}$ とする．以下，このような処理を一定回数 (L) 繰り返した後，出力された画像を既存の bicubic 法などで縮小し，これを次の画素値更新の対象画像 $T_2^{(0)}$ とし，同様の操作を行う．この画素値更新の L 回繰り返しの後，出力画像を縮小するという一連の処理により，Bidirectional Similarity の損失量 d を抑えつつ，出力画像候補のサイズを徐々に小さくしていくことが可能である．この一連の処理を K 回繰り返し，目的のサイズ（ここまでの例では入力画像を 0.95^K 倍縮小したサイズ）になった画像 $T_K^{(L)}$ を最終的な出力画像とする（図 2.4，図 2.5）．

上記のように，Simakov らの手法においては，画素値更新の際に用いる初期値を T の縮小により得ている．この初期値の決定が，単純な画像縮小処理のみで行えるのは，出力画像の形状を矩形領域に限定しているためである．そのため，本手法を任意形に拡張するためには，最終的に任意形の表示領域が得られるように，反復毎の適切な初期値を設定する問題を解く必要があり，そのような手法は自明ではない．

2.4 ワーピングベースアプローチによるイメージリターゲッティング

ワーピングベースは，近年盛んに検討されているアプローチであり，原画像を一定間隔の格子に区切った際の格子点の座標に関する最適化問題を解き，出力画像を得るものである．

ここで，前述のとおりワーピングベースアプローチにおいて出力画像を得る枠組みはすべての手法において共通であり，各手法の差異は格子点の座標の決定問題をどのように解くかという点のみである．この出力画像を得る共通の枠組みの詳細については，本論文の提案手法のひとつである，第 4 章「ワー

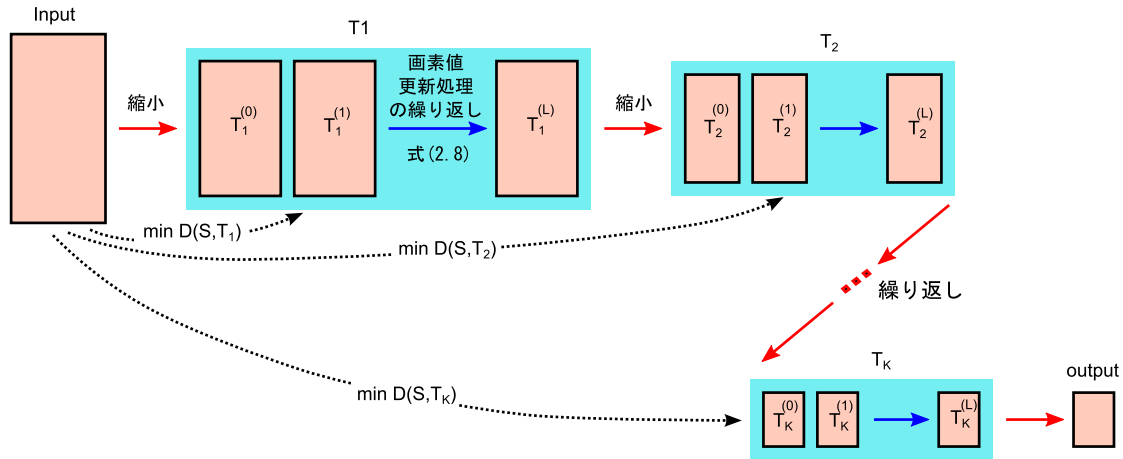


図 2.4 Simakov らの手法 [2] の概要



(a) 原画像

(b) 出力画像（図 2.4 の T_4 に相当）

図 2.5 Simakov らの手法の一例（文献 [2], Figure 5 より引用）

「ワーピングベースの任意形イメージリターゲットング」において述べるため、本節では既存手法に共通する問題点についてのみ述べることにする。

矩形の表示領域を対象としたワーピングベースアプローチの手法 [7, 8, 9, 10, 11, 12, 13, 14] では、各々の文献において定められた凸性をもつエネルギー汎関数上で、大域最適解となる格子点の座標を導出している。これらの手法では、原画像と矩形表示領域の画像境界の形状の対応関係を利用することで、原画像の境界上の格子点を、表示領域の境界に固定するという制約の設定を可能にしている。しかしながら、この制約を任意形の表示領域に拡張する場合には、原画像の境界上の格子点をどの座標に固定するべきか、という問題を解く必要があり、これは自明ではない。

一方で、Feature-Aware Texturing[3] と呼ばれるワーピングベースの手法は、画像を任意形の領域で表示することを可能としており、矩形の表示領域を対



図 2.6 Feature-aware texturing の一例（文献 [3], p. 230, Figure 4 より引用）

象とした手法 [7, 8, 9, 10, 11, 12, 13, 14] と同様に，出力される格子点は全域最適解として与えられる．そのため，一見すると Feature-Aware Texturing は提案手法の目的を解決する手法であるように思われる．しかしながらこの手法は，入力画像に含まれる ROI の形状を保ちつつ，3D モデルに貼り付けることを目的としており，入力画像の画素値はすべて表示領域内にマッピングされる．そのため ROI 以外の領域に関しては，表示領域の境界の形状の影響を大きく受けるため，図 2.6(b) の水平線のように表示領域の境界に沿って画像が大きく歪む結果となる．

これに対して提案手法の目的は，任意形の表示領域において「ROI の表示を可能とする」ことと「全体の歪みを抑えた画像を生成する」ことであり，このような歪みは望ましくない．そのため，ROI の表示領域内への移動を行いながら，背景に関しては原画像のレイアウトを保持するといった問題を解く必要がある．

以上の議論から，全域最適性をもった任意形イメージリターゲッティングを実現するためには，ワーピングベースアプローチを基とし，表示領域の形状に合わせた ROI の移動問題を凸最適化問題として定式化する必要がある．

2.5 むすび

本章では，既存のイメージリターゲッティングに対して，そのアプローチによる分類を行いそれぞれの代表例について述べるとともに，任意形の表示領

域を対象とした際の各々の問題点について言及した。

ピクセルベースに分類される手法は、それぞれ独自のアルゴリズムを提案しているものの、いずれにおいてもアルゴリズムの設計に、表示領域が矩形であるという制約を手がかりとして利用していることがわかる。また、ワーピングベースに分類される手法はすべて、格子点の座標を決定するという共通する枠組みを採用しており、いずれも表示領域の境界に格子点を固定するという制約が必要であるため、単純に任意形の表示領域を対象として手法を拡張することは困難であることを示した。

次章以降では、それぞれの分類における任意形イメージリターゲッティングの詳細について述べる。

第3章

提案手法1：ピクセルベースの任意形イメージリターゲッティング

3.1 まえがき

本章では、前章で述べた既存研究 [2] における画像評価指標 Bidirectional Similarity (Completeness と Completeness) に基づいた、ピクセルベースの任意形イメージリターゲッティング (提案手法1) を提案する。

提案手法1では、矩形画像を任意形ディスプレイ全体を使用するように表示した場合 (図 1.3(a))、損なわれる情報は Completeness であることに着目し、まず Completeness に大きく寄与する画像領域 (矩形とは限らない) を抽出し、それらの領域に対して、Completeness, Coherence を保存するように設計されたエネルギー汎関数に基づく領域移動を行う。その後、移動した画像領域を初期値として Coherence を保存するような画像を生成することで、Bidirectional Similarity を保存しつつ、対象画像の形状によらない画像生成を可能とする。つまり、提案手法は画像変換処理を

1. 画像内の重要性の高い領域の抽出と移動
2. 重要性の低い領域の画素値の再構成

の2段階に分けるアプローチをとる。なお、抽出された重要領域の移動では、既存研究 [2] における画像評価指標や、画像内の重要性の高い領域間の位置関係などを考慮したエネルギー汎関数を新たに定義し、これに対して貪欲法を用いた最小化による移動を行う。また、提案手法1によって変換した後の任意

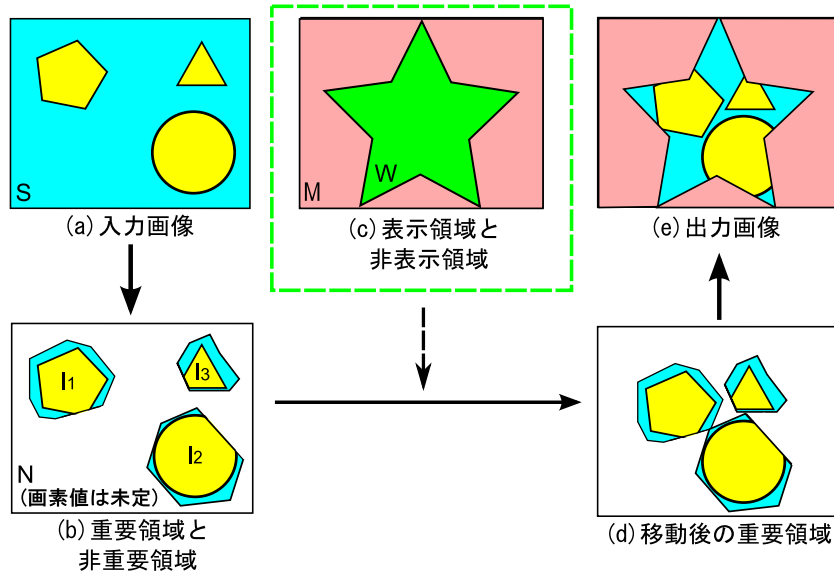


図 3.1 提案手法 1 の概要

形画像に対し、客観指標 Bidirectional Similarity による評価を行うことで、本研究の有効性を示す。

3.2 提案手法 1

図 3.1 に提案手法 1 の概要を示す。本章では、任意形の表示領域を W 、その他の領域を「非表示領域 (M)」とする (図 3.1(c))。また、表示不可となった場合に、画像内において Completeness を大きく損なうような領域を「重要領域 (I_i)」、その他の領域を「非重要領域 (N)」と呼ぶ (図 3.1(b))。ここで、非重要領域の画素値は変換の過程 (図 3.1(b)→(d)) においては未定である。いま、入力矩形画像全体の領域を S とすると、これらの領域の関係は次式のように表すことができる。

$$\begin{aligned}
 S &= \bigcup_i I_i \cup N \\
 &= W \cup M
 \end{aligned} \tag{3.1}$$

提案手法 1 では、重要領域が表示領域に存在するようにすることで Completeness を確保し、その後、重要領域どうしの相対的な位置関係などの画像構造を考慮しつつ画像変換を行うことで Coherence を確保する。具体的には、

以下の三つのプロセスから成る．次節以降において，各プロセスの詳細について述べる．

1. 画像中の重要領域の抽出（Completeness 保存：図 3.1(a)→(b)）
2. 表示領域に応じた重要領域の移動（Completeness, Coherence 保存：図 3.1(b)→(d)）
3. 重要領域を移動させた画像を初期推定とした任意形画像の生成（Coherence 保存：図 3.1(d)→(e)）

3.2.1 画像中の重要領域の抽出

画像中で，非表示領域によって表示不可となった際に Completeness を大きく損なう領域を重要領域として抽出する．いま，Completeness の損失量 d_{comp} は式 (2.5) のように求められる．これは，原画像中において他に似たパッチが存在しないパッチが，出力画像に含まれない（表示できない）場合に，画像中の Completeness が大きく損なわれることを意味している．

上記を踏まえ，原画像 S に対してブロックマッチングを行うことによって重要領域を抽出する．まず， S 内のあるパッチ P を選択する．このパッチ P に対して同原画像内のパッチ Q を探索し，閾値 κ を用いて以下のようにパッチ P の類似パッチ数 n_p を求める．

$$n_p = \sum_{Q \in S} f(D(P, Q)) \quad (3.2)$$

$$f(x) = \begin{cases} 1 & (x \leq \kappa) \\ 0 & (x > \kappa) \end{cases} \quad (3.3)$$

原画像内において，重なりを許し一画素ずつずらしたすべてのパッチ A に対しこの操作を行い，原画像の全パッチに対する類似パッチ数 n_a を求める．こうして求められた n_a がしきい値 λ 以下，すなわち「原画像中で他に似ているパッチが少ない」パッチを重要領域とする．今回，パッチは 10×10 [pixel] の矩形領域とし，それぞれのしきい値は発見的に決定した．

3.2.2 表示領域に応じた重要領域の移動

前節において抽出された i 番目の重要領域が含む画素の集合 I_i ($i \in H := \{1, 2, \dots, h\} \mid h$ は、抽出された重要領域の総数) と、その各集合ごとに算出される重心の座標を用いて、領域の移動を行う。ここで、生成する画像は画像内容を表す二つの指標を保存しているものが望ましく、重要領域を表示領域に移動させることによって、Completeness を確保するだけでなく、重要領域の初期位置や相対位置関係といった画像構造を考慮することで、次節の画像の再構成において Coherence を損なわないようにする必要がある。そこで、各重要領域 I_i に対して、以下のようにエネルギー $E(I_i^{(t)})$ を定義する。

$$E(I_i^{(t)}) = \alpha E_h(I_i^{(t)}) + \beta E_c(I_i^{(t)}) + \gamma E_p(I_i^{(t)}) \quad (3.4)$$

$$E_h(I_i^{(t)}) = \frac{|p|}{|I_i^{(t)}|}, (p \in (I_i^{(t)} \cap M)) \quad (3.5)$$

$$E_c(I_i^{(t)}) = 1 - \frac{1}{h-1} \sum_{j \in \hat{H}_i} \frac{(g(I_i^{(t)}) - g(I_j^{(t)})) \cdot (g(I_i^{(0)}) - g(I_j^{(0)}))}{\|g(I_i^{(t)}) - g(I_j^{(t)})\|_2 \|g(I_i^{(0)}) - g(I_j^{(0)})\|_2} \quad (3.6)$$

$$E_p(I_i^{(t)}) = \|g(I_i^{(t)}) - g(I_i^{(0)})\|_2 / \sqrt{x_{\max}^2 + y_{\max}^2} \quad (3.7)$$

ここで α, β, γ は各項に対する重み付け係数であり ($0 \leq \alpha, \beta, \gamma \leq 1$, ただし $\alpha + \beta + \gamma = 1$), $I_i^{(t)}$ は重要領域の移動の過程における反復処理の際に、 t 回目の移動処理が終了したときの各領域を表す。すなわち、 $I_i^{(0)}$ は各重要領域の初期値である。また、 $g()$ は各領域の重心の位置を 2 次元ベクトルとして返す関数であり、 $\|\cdot\|_2$ は、以下のようにベクトル $\mathbf{x} = (x, y)$ の二乗ノルムを返す。

$$\|\mathbf{x}\|_2 = \sqrt{x^2 + y^2} \quad (3.8)$$

また、 $|\cdot|$ は集合に含まれる画素数を返し、 $x_{\max}, y_{\max}$ はそれぞれ原画像幅、高さを表す。 $\hat{H}_i$ はすべての重要領域番号の集合から i のみを除いた集合 ($\hat{H}_i := \{1, 2, \dots, h\} \setminus i$) である。

E_h は非表示領域内の重要領域の大きさを表すエネルギーであり、表示領域中の重要領域が多いほど小さな値を取る。 E_c はある重心とその他の重心がなすベクトルの初期値からの角度の変位を表す値であり、重心どうしを結んだ図形が原画像のそれと相似形に近いほど小さな値を取る。 E_p は重心の初期位置からの変位量であり、初期位置からの移動量が小さいほど小さな値を取る。

そのため、それぞれのエネルギーがより小さくなるような処理を行った場合、 E_h は表示領域中の重要領域が増加するように働き、 E_c , E_p , は重要領域の原画像からの変位や構造の変化を抑制するように働く。これを式 (2.5), (2.6) の評価指標の観点から見ると、 E_h は Completeness を保存するように、 E_c , E_p は Coherence を保存するように働く。

いま、ある重要領域を上下左右の4方向に1ピクセルだけ移動させた領域を返す演算を、それぞれ S_1, S_2, S_3, S_4 とする。また、 S_0 を重要領域を移動させない演算 ($S_0(I_i) = I_i$) とすると、表示領域に応じた重要領域の移動は Algorithm 1 に従って行われる。

Algorithm 1 重要領域の移動アルゴリズム

Require: $d \in \{0, 1, \dots, 4\}$

$t = 0$

$I_i^{(0)} = I_i, \forall i \in H$

repeat

for $i = 1$ to h **do**

$d_i^* = \arg \min_d (E(S_d(I_i^{(t)})))$

$I_i^{(t+1)} = S_{d_i^*}(I_i^{(t)})$

end for

$t = t + 1$

until $d_i^* = 0, \forall i \in H$

ここで、 $\arg \min_x f(x)$ は、 $f(x)$ を最小とする変数 x を返す関数である。

この操作により生成された画像では、表示領域 W において、各重要領域 I_i ($i \in H$) にのみ、それぞれが原画像の該当する位置の画素値がそのまま与えられており、3.2.1 節で画素値を破棄した領域 ($S \setminus \bigcup_i I_i$) については画素値が未定の状態である。以下、このような画像を初期推定画像とよび、次節において、未定となっている画素値を Coherence を保存するように決定する手法について述べる。

3.2.3 初期推定画像を用いた任意形画像の生成

前節までで生成された初期推定画像は、Completeness に大きく寄与する領域が表示領域に移動したものであり、Completeness が確保された状態である。本

節では、初期推定画像において未定となっている画素値を、Coherenceを保つように推定し、最終的な任意形画像を生成する。

いま、Coherenceの損失量 d_{cohere} は式(2.6)のように求められる。そこで、入力画像(S)と前節までで求めた初期推定画像(T_1)を用いて、以下のように未定の画素値を決定する。

まず、重要領域 I_i およびその周囲1画素(値は未定)の領域を T' とする。この T' に含まれる画素 q' の値 $T'(q')$ を、以下の式により更新する。

$$T'(q') = \frac{\sum_{i=1}^w S(p'_i)}{w} \quad (3.9)$$

ここで、画素 q' を含み、かつすべての画素が T' に含まれるパッチ Q'_i に対して、 $D(\cdot)$ が最小となる S 内のパッチ P'_i を考える。この時 p'_i は、パッチ Q'_i における画素 q' の位置と対応するパッチ P'_i 内の位置の画素を表している。また、 w は画素 q' を含み、かつすべての画素が T' に含まれるパッチ Q'_i の総数である。この操作により、初期推定画像より1画素分だけ画素値の確定した領域が広がった画像(T_2)の生成が可能である。さらに、この T_2 において、画素が確定している領域およびその周囲1画素の領域を次の T' とし、同様の操作を行う。この操作の繰り返しにより、初期推定画像において画素値が未定であった領域の画素値を徐々に決定することが可能となる。

以上の操作を、画素値が未定の領域($S \setminus \bigcup_i I_i$)のすべての画素値が決定するまで反復することにより画像の再構成を行い、求める任意形画像を得る。

3.3 実験および考察

本節では、提案手法1の有効性を示すため、既存のピクセルベースのイメージリターゲットング手法との比較実験を行う。実験の際の客観評価指標には、2.3.2節で述べた、Bidirectional Similarityを用いた。また、提案手法1のパラメータ設定に対する頑健性を確認するための実験も行った。

3.3.1 Bidirectional Similarity と主観評価との相関

Rubinsteinらは文献[31]において、既存の様々なイメージリターゲットング手法の主観評価実験を行うとともに、その主観評価値と、いくつかの客観

評価指標との間の順位相関を示した。彼らはまず、37枚の自然画像に対して、既存の8種類の画像変換手法を適用し、その結果画像に対して主観評価実験による順位と、Bidirectional Similarityなどの客観評価指標による順位を付与した。次に、これら二つの順位の間でケンドールの順位相関係数を算出し、これにより主観評価と客観評価指標との間で順位相関の評価を行った。文献[31]では、Bidirectional Similarityは「顔や人物」、「草原やレンガなどの繰り返し性をもったパターン」、「前景にはっきりとしたオブジェクトが存在する」といった要素をもつ画像に対して、主観評価と高い順位相関を示すことが明らかにされている。

3.3.2 比較実験

図3.2¹に、いくつかの原画像に対して、「原画像をそのまま任意形の表示領域で表示した画像」、「既存手法を用いて任意形の表示領域に外接するように変換した画像」、および「本提案手法1を適用した結果画像」を示す。原画像（図3.2(a)）をそのまま任意形の表示領域で表示した画像（図3.2(b)）は、非表示領域に含まれる画素の値を0（黒色）とすることによって表現されており、図3.2(b)の黒色の領域を含んだ画像のサイズは、原画像と同等である。また、既存手法および提案手法1を適用した図3.2(c)～(f)はそれぞれ、図3.2(b)と同サイズの任意形の表示領域に対する結果であり、図3.2(c)のキャプションに記載されたサイズは、表示領域に外接する最小の矩形領域のサイズを示している。この矩形領域のサイズは図3.2(d)～(f)においても同様である。本実験における、既存のイメージリターゲットングとしては、“bicubicによる均一な縮小（図3.2(c)）”、“Seam Carving[1, 4, 5]（図3.2(d)）”、“Image Summarizing via Bidirectional Similarity[2]（以下、ISBS）（図3.2(e)）を用いた。なお、Seam Carvingを実行するうえで必要となる、画素の重要性の指標としては、文献[1, 4, 5]にならい、近隣画素との輝度値差分を用いた。また、ISBSでは、画像サイズを変える $T_k^{(L)}$ から $T_{k+1}^{(0)}$ （ $k \in \{1, \dots, K-1\}$ ）への縮小の際に、いずれの画像においても3～6 [pixel]だけ画像幅が小さくなるように、bicubicを用いて縮小することとし、画素値更新の反復回数は10回（ $L=10$ ）とした。なお、提案

¹The authors are permitted to publish and modify of the photographs released under the Creative Commons. The photographs are from U.S. Fish and Wildlife Service - Midwest Region's photostream and Steve-h.

手法1適用結果については、各パラメータ ($\kappa, \lambda, \alpha, \beta, \gamma$) を良い結果画像が得られるように選んだときの結果を掲載している。使用されたパラメータの値は、それぞれの画像のキャプションに示されている。

これらの結果画像に対して、Bidirectional Similarityを用いた評価実験を行う。まず、入力画像を原画像、出力画像を図3.2(b)~(f)に示した、任意形の表示領域における5種類の表示手法による画像とする。出力画像の入力画像に対する画像内容の損失量 d_{comp} , d_{cohere} , およびその重み付き和 d ($\varepsilon = 0.5$) を計測する。ここで入力画像、出力画像のそれぞれのパッチは 10×10 [pixel] の矩形領域とし、探索範囲はパッチに含まれる画素すべてが有効である（表示領域内となる）領域とした。表3.1~3.4にそれぞれの画像に対する d_{comp} , d_{cohere} , および d を示す。ここで、式(2.5)~(2.7)の d_{comp} , d_{cohere} , d は、それぞれ Completeness, Coherence, Bidirectional Similarity の損失量を表しているため、この値が大きいほど二つの画像は似ていないことを意味していることに注意されたい。

表3.1~3.4をみると、画像 dolphin に対して Seam Carving を適用した場合を除いたすべての画像に対して、提案手法1適用結果の方が小さい d の値を示しており、より画像全体の Bidirectional Similarity を保存できていることがわかる。また、画像 dolphin に対して Seam Carving を適用した場合においても、提案手法1の適用結果はほぼ同等の d の値となっており、Bidirectional Similarity の損失量に大きな差はみられない。以上より、本手法が客観評価指標の上で有効であることが示された。

3.3.3 パラメータ設定における頑健性に関する考察

図3.3は、それぞれの原画像に対して、本手法により重要領域の抽出を行い、異なるパラメータにおける重要領域の移動の結果を示した画像の一例である。各画像のパラメータはそれぞれのキャプションを参照されたい。表3.5は、移動の際のパラメータ (α, β, γ) を、 $\alpha + \beta + \gamma = 1.0$, $\alpha, \beta, \gamma > 0$ の条件のもとで0.1ずつ変化させ、それらに対応する提案手法1適用結果画像に対して、Bidirectional Similarity を算出し、表にしたものである。図3.3から、 E_h, E_c, E_p の項、それぞれに対応するパラメータが3.2.2節で述べた性質を保つよう適切に作用していることがわかる。また、表3.5から、パラメータ (α, β, γ) の値の変化によって、Bidirectional Similarity の損失量 d は極端に変化していないこと、ほぼすべてのパラメータにおいて、提案手法1による結果は表3.2の既存手法

表 3.1 評価実験結果 1 (bison [250 × 188 pixel])

bison [250 × 188 pixel]	d_{comp}	d_{cohere}	d
原画像を任意形の表示領域で表示	45.91	0	22.96
bicubic	45.32	8.20	26.76
SeamCarving	47.62	7.01	27.32
ISBS	41.55	2.42	21.99
提案手法 1	40.79	1.11	20.95

による結果の d を下回っていることがわかる．以上の結果より，本手法はパラメータの設定に対して一定の頑健性をもっているといえる．

3.3.4 考察

比較実験により，本手法が客観評価指標の上で有効であることを示し，パラメータ設定に対する頑健性も確認した．次に，図 3.2 の各結果に対する考察を行う．まず，原画像をそのまま任意形の表示領域で表示した場合や，原画像 fish に対して Seam Carving を適用した場合などは，視覚的に重要なオブジェクトが非表示領域により表示不可能となっていることがわかる．また，bicubic を用いて縮小した場合は，オブジェクトが全体的に縦長となってしまう，形状を損なっているといえる．ISBS の場合も，bicubic による縮小の影響をうけるため，同様にオブジェクトの形状を損なっている．さらに，原画像 human-bird に関しては，既存手法いずれにおいても，左下の人物の全体が表示されておらず，またその形状も損なっている．これは既存手法は画像幅を縮小する場合，画像内のすべての領域に対して，水平方向へ縮小する効果しかもたないことに起因する．これに対し，提案手法 1 適用結果では，魚や人などの視覚的に重要なオブジェクトを，形状を保存しつつユーザーに提示可能となっていることがわかる．

表 3.2 評価実験結果 2 (fish [320 × 240 pixel])

fish [320 × 240 pixel]	d_{comp}	d_{cohere}	d
原画像を任意形の表示領域で表示	47.21	0	23.61
bicubic	55.43	52.12	53.76
Seam Carving	52.69	23.04	37.87
ISBS	47.11	28.38	37.75
提案手法 1	38.54	3.09	20.82

表 3.3 評価実験結果 3 (dolphin [320 × 240 pixel])

dolphin [320 × 240 pixel]	d_{comp}	d_{cohere}	d
原画像を任意形の表示領域で表示	6.20	0	3.10
bicubic	5.18	4.69	4.92
Seam Carving	3.25	0.39	1.82
ISBS	3.81	7.79	5.8
提案手法 1	3.52	0.15	1.84

表 3.4 評価実験結果 4 (human-bird [300 × 225 pixel])

human-bird [300 × 225 pixel]	d_{comp}	d_{cohere}	d
原画像を任意形の表示領域で表示	241.57	0	120.78
bicubic	83.45	28.24	55.84
Seam Carving	76.49	15.55	46.02
ISBS	70.99	8.21	39.60
提案手法 1	48.65	7.25	27.95

表 3.5 異なるパラメータ (α, β, γ) における Bidirectional Similarity の損失量 d .
 $(\gamma = 1.0 - (\alpha + \beta))$

$\alpha \backslash \beta$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.1	22.78	21.86	21.89	22.80	23.19	22.91	22.87	22.93
0.2	23.09	22.11	22.11	23.17	23.08	23.17	22.93	-
0.3	23.26	20.81	22.50	22.74	23.23	23.09	-	-
0.4	22.28	20.72	23.15	23.13	23.23	-	-	-
0.5	22.00	22.74	22.57	23.23	-	-	-	-
0.6	21.95	22.48	23.79	-	-	-	-	-
0.7	20.63	20.82	-	-	-	-	-	-
0.8	20.57	-	-	-	-	-	-	-

3.4 むすび

本章では，提案手法1として，ピクセルベースの任意形イメージリターゲッティングを提案した．

提案手法1は，Simakov らの手法 [2] において定義された画像内容を表す客観指標 Completeness, Coherence が画像中のユニークな画像特徴をより良く表現できることに着目し，その両者を保つような画像生成を行っている．Simakov らの手法は，画素値の更新の際に参照する画像を直前の出力画像の縮小処理によって得ているため，これを単純に任意形表示領域上での表示に適用することは不可能であった．これに対し提案手法1では，問題を，Completeness を考慮した重要領域の抽出と移動，Coherence を考慮した非重要領域の画像再構成の2段階に分割するというアプローチにより解決した．抽出された重要領域の移動の際には，Completeness, Coherence 両者を保存するようなエネルギー汎関数を用い，この汎関数が局所最小となるように移動を行った．以上の処理により，二つの客観指標を保った任意形画像を生成することに成功している．

さらに，提案手法1の結果に対して，Completeness, Coherence を用いた評価実験を行なった．4種類の原画像に対する実験結果より，提案手法1がCompletenessを保存しつつ，Coherenceを損なわないように働いていることがわかり，また，視覚的にも重要な領域を含みつつ違和感のない任意形画像を生成できていることが確認できた．

前章において，ピクセルベースアプローチは非凸な問題を解く必要があるため，定められたエネルギー汎関数における大域最適性が保証されないことを述べた．提案手法1についても，重要領域の抽出と移動や非重要領域の画像再構成といった，それぞれのプロセスにおいて局所最小解を導出するアルゴリズムを提案しており，結果として最終的な画像出力に対する何らかの評価基準における大域最適性は保証されない．

この問題を解決するために，次章では，最適化問題と親和性の高いワーピングベースの枠組みに基づいた，任意形イメージリターゲッティングを提案する．

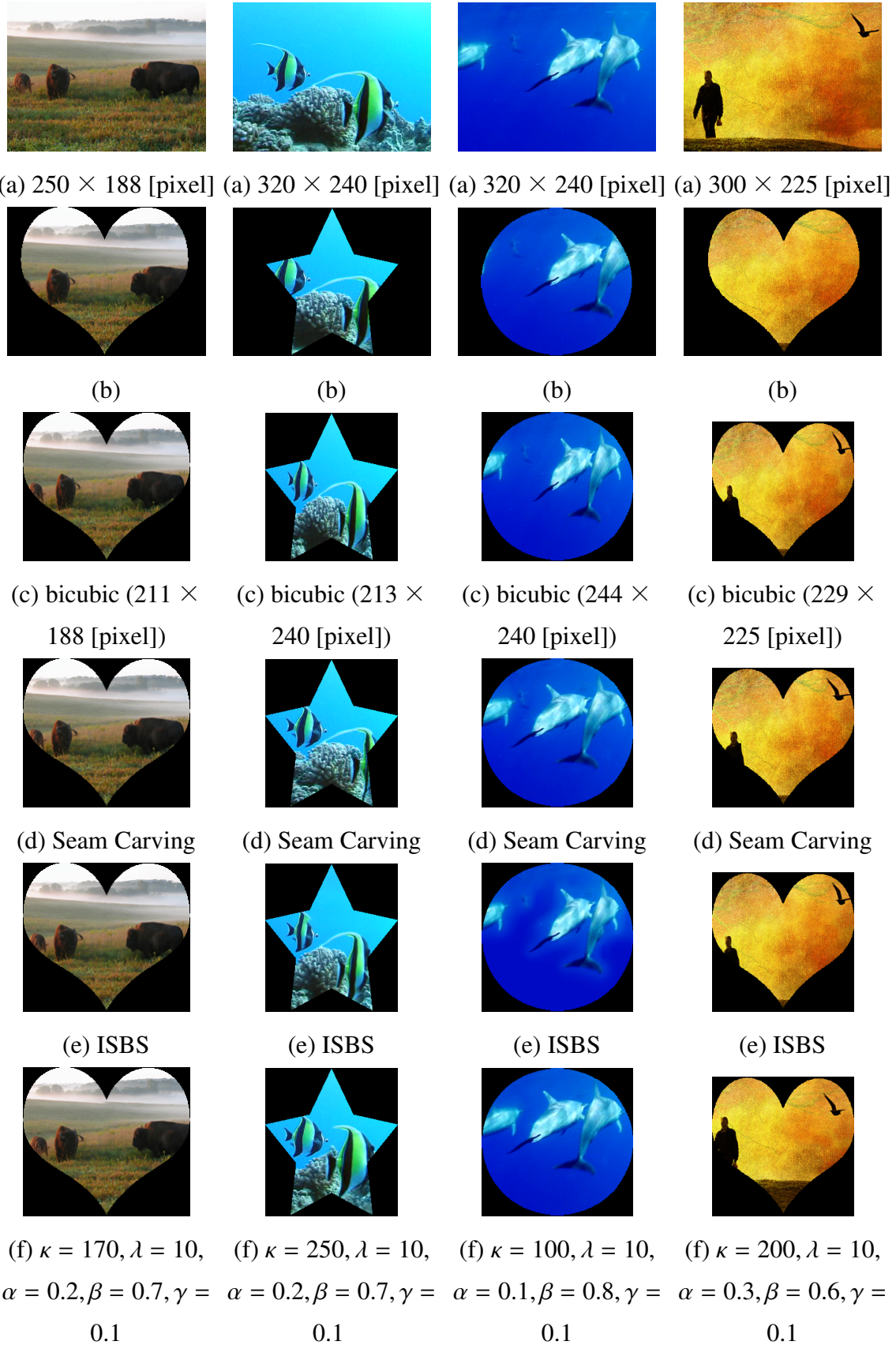


図 3.2 (a) 原画像, (b) 原画像を任意形の表示領域で表示した画像, (c)~(e) 既存手法を用いて, 任意形の表示領域に外接するように変換した画像, (f) 提案手法 1 結果画像, (左から右に [bison, fish, dolphin, human-bird], 黒色: 非表示領域)

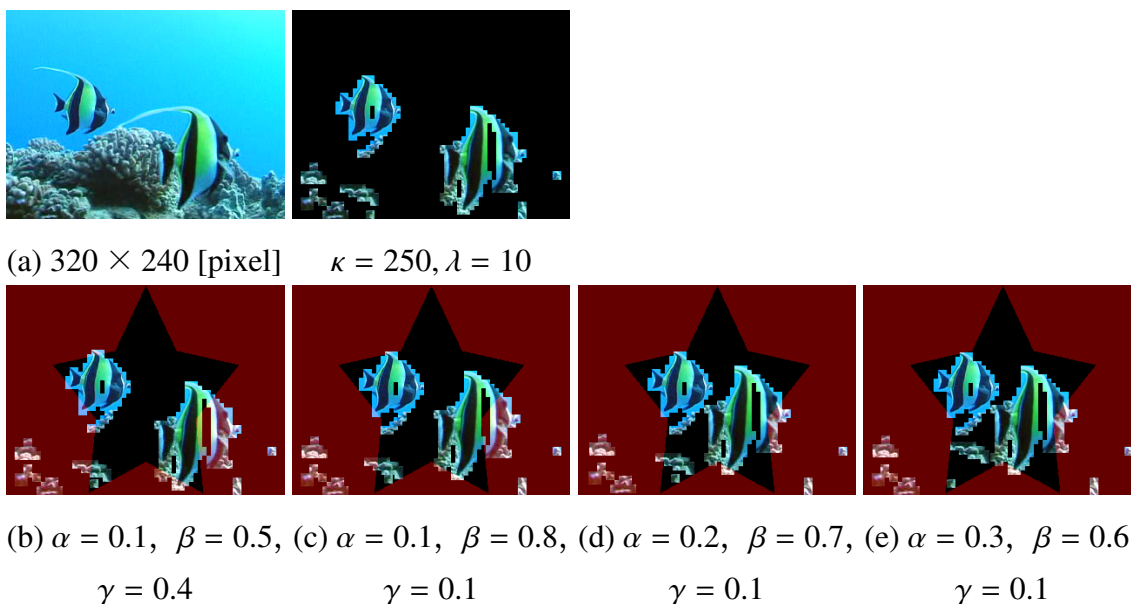


図 3.3 (a) 原画像, (b) 重要領域抽出結果, (c)～(f) 重要領域移動結果, 黒色：値が未定の画素, 赤色：非表示領域, 各画像（黒色, 赤色領域を含む）はいずれも原画像と同サイズ

第4章

提案手法2：ワーピングベースの任意形イメージリターゲッティング

4.1 まえがき

本章では、「任意形の表示領域の形状に合わせたROIの移動」と「歪みのない画像生成」、ならびに「定められた基準における大域最適性の保証」を同時に実現するために、凸最適化を用いたROI移動の大域最適化による任意形イメージリターゲッティングを提案する。

提案手法2のキーアイデアは、原画像における全ROIの座標から各ROIの移動可能領域をあらかじめ算出し、各ROIの移動範囲を当該領域内に制限したことである。これにより、表示領域の形状にあわせたROIの移動問題を凸関数として定式化し、この問題の大域的最適解として、ROIの移動先座標を得ることが可能となる。導出されたROIの座標をもとに、ワーピングベースと呼ばれる既存のイメージリターゲッティングの枠組みに従い、画像全体を再構成することで出力画像が得られる。

4.2 準備

定義 4.2.1 (ノルム) N 実次元ベクトル空間 $\mathbb{R}^N$ において、ベクトル $\mathbf{x} \in \mathbb{R}^N$ の ℓ_2 ノルムを以下のように定義する。

$$\|\mathbf{x}\|_2 := \sqrt{\sum_{i=1}^N (x_i)^2}. \quad (4.1)$$

ここで, x_i はベクトル $\mathbf{x}$ の i 番目の成分である.

定義 4.2.2 (指示関数) ある集合 H の指示関数 $\iota_H(\mathbf{x})$ は以下のように定義される.

$$\iota_H(\mathbf{x}) := \begin{cases} 0, & \text{if } \mathbf{x} \in H, \\ \infty, & \text{otherwise.} \end{cases} \quad (4.2)$$

定義 4.2.3 (下半連続で proper な凸関数) 関数 $f: \mathbb{R}^N \rightarrow \mathbb{R} \cup \{\infty\}$ が以下の条件を満たしているとき, 関数 f は下半連続で *proper* な凸関数であるという.

1. $\text{dom}(f) := \{\mathbf{x} \in \mathbb{R}^N \mid f(\mathbf{x}) < \infty\} \neq \emptyset$.
2. $\text{lev}_{\leq \alpha}(f) := \{\mathbf{x} \in \mathbb{R}^N \mid f(\mathbf{x}) \leq \alpha\}$ が任意の $\alpha \in \mathbb{R}$ に対して閉集合である.
3. $f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}), \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^N, \forall \lambda \in (0, 1)$.

ここで $\emptyset$ は空集合を表す. さらに, $\mathbb{R}^N$ 上のすべての下半連続で *proper* な凸関数の集合を $\Gamma_0(\mathbb{R}^N)$ で表す.

定義 4.2.4 (近接作用素) インデックス $\gamma \in (0, +\infty)$ をもつ関数 $f \in \Gamma_0(\mathbb{R}^N)$ の近接作用素を以下のように定義する.

$$\text{prox}_{\gamma f}: \mathbb{R}^N \rightarrow \mathbb{R}^N: \mathbf{x} \mapsto \arg \min_{\mathbf{y} \in \mathbb{R}^N} \left(f(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{y}\|_2^2 \right). \quad (4.3)$$

定義 4.2.5 (相対的内部) 集合 $H \subset \mathbb{R}^N$ に対して, H を含む最小のアフィン集合を $\text{aff}H$ とする. このとき, H の相対的内部 $\text{ri}(H)$ は以下のように定義される.

$$\text{ri}(H) = \{\mathbf{x} \in H \mid \text{Ball}(\mathbf{x}, r) \cap \text{aff}H \subseteq H, \exists r > 0\}. \quad (4.4)$$

ここで, $\text{Ball}(\mathbf{x}, r) = \{\mathbf{y} \mid \|\mathbf{y} - \mathbf{x}\|_2 \leq r\}$ である.

4.3 提案手法2

4.3.1 概要

本提案手法2は, 既存のワーピングベースの手法 [3, 7, 8, 9, 10, 11, 12, 13, 14] の枠組みを応用し, ひとまず任意形の表示領域 $\tilde{T}$ に外接する矩形領域 T に対して入力画像 S の画素値をマッピングした矩形画像を生成する. そして, そ

の矩形画像に対して $\tilde{T}$ でマスクを行い、 $\tilde{T}$ 外の画素の値を破棄することで最終的な任意形の表示領域での画像表示を行う。このマスクの際に ROI の画素値が破棄されないよう、矩形画像の生成段階において、各 ROI が可能な限り表示領域内に収まるように移動を行う必要がある。

提案手法2の処理は、以下の4つの手順に分割することができる。その概要を図4.1に示す。

1. 原画像内の ROI の抽出
2. 原画像の格子状分割、および表示領域の形状を用いた格子点の座標決定
3. 決定された格子座標に対するテクスチャマッピング（矩形画像出力）
4. 出力された矩形画像の任意形の表示領域によるマスク

ここで、第2章でも述べたように、手順(2),(3)によって矩形画像を得るアルゴリズムの枠組みは、既存のワーピングベースの手法すべてに共通しており、提案手法2では格子点の座標の決定方法のみが異なることに注意されたい。本節では、まずその共通の枠組みについて述べる。

いま、入力画像 S を一定間隔の格子に区切った際に得られる、ある格子点の座標を $v_i := (v_i^x, v_i^y)$, ($i \in \{1, 2, \dots, n\}$) とする。また、これら格子点によって囲まれる最小の画像領域（格子面）を s_j ($j \in \{1, 2, \dots, n_l\}$) とする。ここで、 n , n_l はそれぞれ S に含まれる格子点、および格子面の総数である。この v_i に対応する出力画像 T 内の格子点の座標 u_i を決定すると、同時に対応する格子面 t_j の位置が得られる。さらに、 s_j の画素値を入力として、既存のテクスチャマッピング技術 [32] を用いることで t_j の画素値が得られる。同様の操作を出力画像 T に含まれるすべての t_j に対して行い、矩形の出力画像 T を得る。以上の枠組みは全てのワーピングベースの手法において共通であり、それらの手法の差異を決定する最も重要な要素は、格子点の座標の決定法（最小化するエネルギー汎関数の選択）の違いにあるといえる。本提案手法2では、任意形の表示領域に応じた ROI の移動、および原画像との整合性の保存を実現するエネルギーを凸最適化問題として定式化する。

次節では、手順(1)にあたる ROI の抽出について、4.3.3 節では、手順(2)において格子点の座標を決定する際のエネルギー汎関数の詳細について述べる。

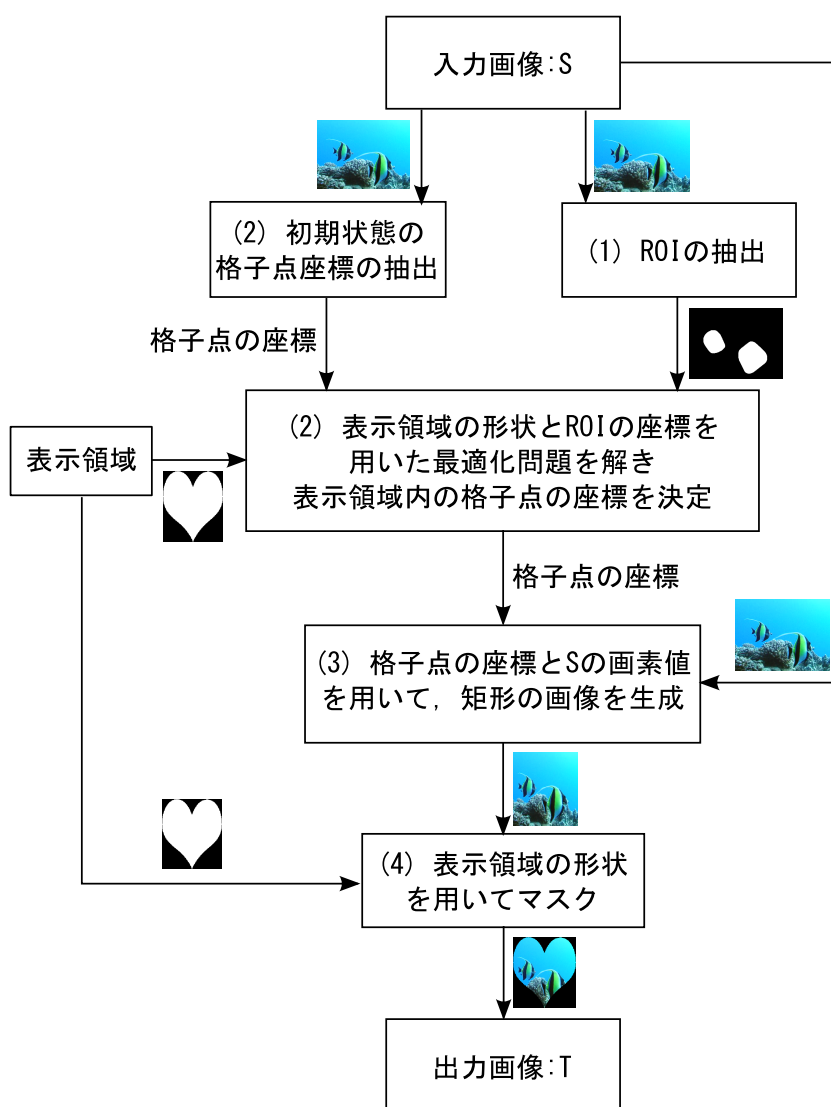


図 4.1 提案手法 2 の概要

4.3.2 ROIの抽出

提案手法2においても多くの他の手法と同様に、まず原画像内のROIを決定する必要がある．2.2節で述べたように、イメージリターゲティングの分野ではROIの決定に画像内の顕著性マップを決定する手法[21, 22, 23, 24, 25, 26]が広く用いられている．本論文においては、このROIの決定に、既存手法のうちのstate-of-the-artの一つとされているHouらの顕著性マップ決定手法[26]を採用した．

図4.2(b)にHouらの手法によって決定された顕著性マップを示す．この画像において、画素値は当該画素の顕著性を表しており、明るい画素ほどより顕著である．得られた顕著性マップに基づき、画像を（一般に複数の）ROIと、それ以外の領域とに分割する．この分割のために、まず閾値 κ を用いた閾値処理を行い、Houらの手法による顕著性マップの値が κ より高い画素をROI候補画素とする（図4.2(c)）．さらに、このROI候補画素を含む領域（図4.2(c)の白色の領域）に対して、その凸包を計算し、画像内のROIとする（図4.2(d)）．

このように凸包を用いてROIの形状を凸形状にすることにより、次節以降で示すエネルギー汎関数の最小化問題を凸最適化問題として定式化することが可能となる．

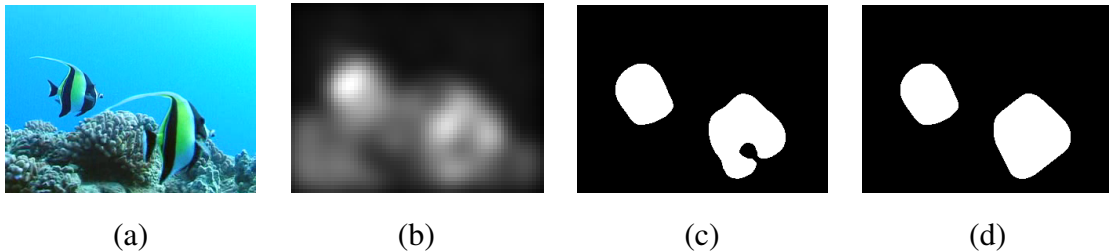


図 4.2 ROI 抽出の一例 (a) 原画像, (b) 顕著性マップ, (c) ROI 候補画素（白色）, (d) ROI

4.3.3 格子点に対するエネルギー汎関数

いま $\mathbf{v} := (v_1, v_2, \dots, v_n)^T \in \mathbb{R}^{2n}$ を、 S を区切った際の初期状態の格子点座標を並べた列ベクトルとし、 $\mathbf{u} := (u_1, u_2, \dots, u_n)^T \in \mathbb{R}^{2n}$ をそれら格子点に対応する T 内の格子点の座標を表す変数ベクトルとする．この $\mathbf{u}$ に対して以下の最適

化問題を定式化する.

$$\begin{aligned} \mathbf{u}^* &= \arg \min_{\mathbf{u} \in \mathbb{R}^{2n}} \zeta(\mathbf{u}) \\ &= \arg \min_{\mathbf{u} \in \mathbb{R}^{2n}} E_{pos}(\mathbf{u}) + E_{shape}(\mathbf{u}) + E_{cohere}(\mathbf{u}). \end{aligned} \quad (4.5)$$

ここで, 3つの項 E_{pos} , E_{shape} , および E_{cohere} はそれぞれ, 「ROI の表示」, 「ROI の形状保存」, および 「原画像との整合性の保存」 のために働くエネルギーとなっている. また, 定式化におけるすべての項に共通する方針として, パラメータチューニングの容易さの観点から可能な限り指示関数を用いている. これは, 指示関数は0または ∞ を出力する関数であり, 指示関数によって構成される項については重みを決定する係数をかける必要がないためである.

式(4.5)に示される3つの項はすべて凸関数であり, 式(4.5)に対して既存の凸最適化問題に対するアルゴリズムを適用することで, 大域最適解 $\mathbf{u}^*$ を得ることが可能である. 以下の節では, 各項の詳細について述べる.

ROIを可能な限り表示するためのエネルギー

4.3.2節において決定されたROIは, 画像内に複数存在し, それらの形状は凸となっている. いま, i 番目のROIが含む画素の座標の集合を I_i とする. さらに, この I_i の座標を表す代表点の座標を $\mathbf{g}_i \in \mathbb{R}^2$ とする. 初期状態において $\mathbf{g}_i \in \mathbb{R}^2$ は各 I_i の重心の座標に最も近い格子点の座標である. この $\mathbf{g}_i \in \mathbb{R}^2$ で表される格子点を, 以下では (i 番目のROIの) 最近傍格子点と呼ぶ. すなわち最近傍格子点の総数 k はROIの総数と等しい. この最近傍格子点の座標について, 以下のように凸関数を定義する.

$$E_{pos}(\mathbf{u}) := \sum_{i \in K} \frac{\lambda_1}{2} (d_{A_i}(R_i \mathbf{u}))^2 + \frac{\lambda_2}{2} \|\mathbf{G}\mathbf{u} - \mathbf{G}\mathbf{v}\|_2^2 + \sum_{i \in K} \sum_{j \in K \setminus i} \iota_{W_{ij}}(R_i \mathbf{u}), \quad (4.6)$$

ここで, $R_i \in \mathbb{R}^{2 \times 2n}$ は i 番目の最近傍格子点の座標を $\mathbf{u}$ から出力する行列であり, $K := \{1, 2, \dots, k\}$ はROIの番号を指定するインデックスの集合である. 距離関数 $d_C(\mathbf{x}) := \|\mathbf{x} - P_C(\mathbf{x})\|_2$ はある点 $\mathbf{x}$ からある凸集合 C への距離を表し, $P_C: \mathbb{R}^2 \rightarrow \mathbb{R}^2: \mathbf{x} \mapsto \arg \min_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|_2$ は $\mathbf{x}$ から C への距離射影である. さらに, $\mathbf{G} \in \mathbb{R}^{2k \times 2n}$ は $\mathbf{u}$ からすべての最近傍格子点の座標をならべた列ベクトルを出力する行列である.

式(4.6)第1項における A_i は, 表示領域 $\tilde{T}$ において i 番目のROIがすべて表示されるための, 最近傍格子点の取りうる座標集合を表しており, 以下のよ

うに定義される (図 4.3) .

$$A_i := \{\mathbf{p} \mid I_i \mathbf{p} \subseteq \hat{T}\}, \quad (4.7)$$

ここで, $\hat{T}$ は $\tilde{T}$ の凸包であり, $I_i \mathbf{p}$ はベクトル $\mathbf{p}$ により I_i を移動させた先の座標集合を表す. 式 (4.7) は, 初期状態における $\mathbf{g}_i$ の座標を原点として, ミンコフスキー差 $\ominus$ を用いて以下のように表すことができる.

$$A_i = \hat{T} \ominus I_i. \quad (4.8)$$

ただし,

定義 4.3.1 (ミンコフスキー差) 実ベクトル空間 $\mathbb{R}^N$ の二つの部分集合 $P, Q \in \mathbb{R}^N$ に関して, ミンコフスキー差 $\ominus$ は以下のように定義される.

$$P \ominus Q := \bigcap_{\mathbf{q} \in Q} \{\mathbf{p} - \mathbf{q} \mid \mathbf{p} \in P\}, \quad (4.9)$$

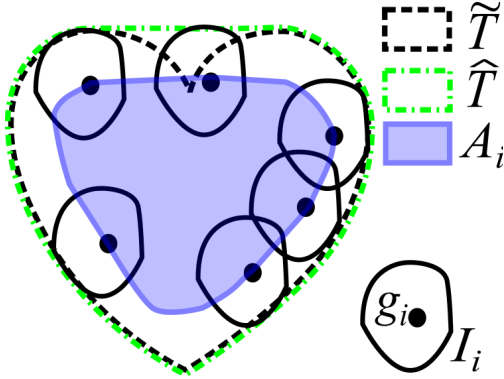
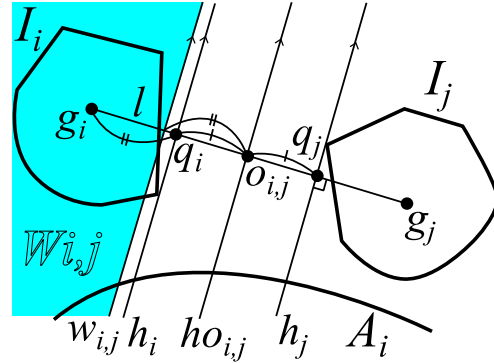
いま, 4.3.2 節の処理により, すべての ROI は凸形状であり, さらに $\hat{T}$ も同様に凸形状である. そのため, 式 (4.8) で定義される A_i は凸形状をもつ (文献 [33, 34, 35] 参照). これにより, 各 ROI から A_i に対しての距離関数を, 距離射影を用いて計算することが可能となる.

また, i 番目の ROI が A_i に近づくほど $(d_{A_i}(R_i \mathbf{u}))^2$ の値は小さくなる. そのため, 式 (4.6) 第 1 項は表示領域内に ROI を収めるように働くエネルギーとなる. 一方, 式 (4.6) 第 2 項は ROI の座標を初期状態の座標 (原画像での座標) に近づける働きをもつ. ROI が初期座標から大幅に移動した場合, 結果として出力画像におけるアーティファクトを招く恐れがあり, それを抑制するためにこの式 (4.6) 第 2 項が必要となる. つまり, これら 2 項は, すべての ROI が表示可能となるように表示領域に移動させようとするエネルギーと, 結果画像のアーティファクトを抑制するために初期座標に固定しようとするエネルギーであり, 互いにトレードオフの関係になっている.

本研究の目的を実現するためには, 基本的にすべての ROI が表示領域に配置されることが望ましい. しかしながら, 式 (4.6) 第 1 項, 第 2 項のみによる移動を行った場合, 表示領域や ROI の形状により, ROI 同士のオーバーラップが発生する場合がある. これを避けるために, 提案手法 2 では式 (4.6) 第 3 項を定義し, 各 ROI の移動可能領域を決定している. 式 (4.6) における $W_{i,j}$ は i 番目と j 番目の ROI の間を通る直線を境界とする半空間を表しており, これが各 ROI の移動可能領域を決定している (図 4.4) .

まず、初期状態において、その最近傍格子点 g_i と g_j を通る線分 l を引くことができる。この l に垂直に交わり、 i 番目の ROI に接する直線を h_i とし、その交点を q_i とする (j 番目の ROI に対しても同様に、 h_j , q_j が算出される)。さらに、 q_i と q_j の中点 $o_{i,j}$ を通り、 l に垂直な直線を $ho_{i,j}$ とする。この $ho_{i,j}$ は、 i 番目の ROI と j 番目の ROI とが互いに等しい速さで近づきあった場合でも、決してオーバーラップしない領域を決定する境界線である。実際には、移動後の i 番目の ROI の座標は最近傍格子点で表わされるため、 g_i と q_i の距離だけ $ho_{i,j}$ を平行移動させた直線 $w_{i,j}$ が、 i 番目の ROI の j 番目の ROI に対する移動可能領域 $W_{i,j}$ の境界線となる (図 4.4)。

すべての $W_{i,j}$ は初期状態で算出されるものであり、最適化問題を解く際の各反復処理の中では更新されないことに注意されたい。ここで、もし $w_{i,j}$ と A_i とが交わりをもたない場合は、 $w_{i,j}$ を平行移動させた際に、新たに得られる $W_{i,j}$ と A_i が交点をもち、かつ $W_{i,j}$ の面積が最小となる位置へ $w_{i,j}$ を平行移動し、これを $w_{i,j}$ として再定義する。

図 4.3 A_i の概略図 4.4 $W_{i,j}$ の概略

ROI の形状保存のためのエネルギー

4.3.2 節において決定された ROI は、視覚的に顕著な画像領域であるため、その歪みや不自然なアーティファクトは可能な限り避ける必要がある。式 (4.5) の第 2 項、 $E_{shape}(\mathbf{u})$ は ROI の形状を保つように働くエネルギーであり、以下のように定義される。

$$E_{shape}(\mathbf{u}) := \iota_{H_f}(F\mathbf{u}), \quad (4.10)$$

$$H_f := \{\mathbf{x} \mid \mathbf{x} = F\mathbf{v}\}. \quad (4.11)$$

ここで, $F \in \mathbb{R}^{4n \times 2n}$ は, ROI に含まれるすべての格子点について, その右方向と下方向に隣接する格子点との座標の差分を $\mathbf{u}$ から算出する行列である. つまり, H_f は初期状態の格子点座標に F を作用させた際に得られる差分のベクトルそのものであり, これは明らかに凸性をもつ.

原画像との整合性を保つためのエネルギー

4.3.3 節では ROI の歪みを抑制するエネルギーを定式化したが, 同様に ROI でない画像領域についても出力画像において可能な限り歪みを抑えるべきである. 以下に示す式 (4.5) の第3項 $E_{cohere}(\mathbf{u})$ は, この歪みを低減するエネルギーである.

$$E_{cohere}(\mathbf{u}) := \frac{\lambda_3}{2} \|D\mathbf{u}\|_2^2 + \iota_{H_b}(B\mathbf{u}) + \iota_{H_l}(\mathbf{u}), \quad (4.12)$$

ここで, $D \in \mathbb{R}^{4n \times 2n}$ は画像に含まれるすべての格子点について, その右方向と下方向に隣接する格子点との座標の差分を算出する行列である (4.3.3 節における F と類似しているが, D は ROI を含む全ての画像領域に含まれる格子点を出力することに注意されたい).

また, $\iota_{H_b}(B\mathbf{u})$ は原画像 S において画像境界に含まれる格子点を, 表示領域に外接する矩形領域 T の画像境界に固定するものであり, $\iota_{H_l}(\mathbf{u})$ は以下のように定義される.

$$\begin{aligned} H_b &:= \{\mathbf{x} \mid \mathbf{x} = B\mathbf{u}, B_{lx}\mathbf{u} = \mathbf{1}_{(n_v)}, B_{rx}\mathbf{u} = w_T \mathbf{1}_{(n_v)}, \\ &\quad B_{ty}\mathbf{u} = \mathbf{1}_{(n_h)}, B_{by}\mathbf{u} = h_T \mathbf{1}_{(n_h)}, \\ &\quad B = (B_{lx}^T B_{rx}^T B_{ty}^T B_{by}^T)^T\}. \end{aligned} \quad (4.13)$$

$$H_l := \{\mathbf{u} \mid 1 \leq \mathbf{u}_i^x \leq w_T, 1 \leq \mathbf{u}_i^y \leq h_T, i \in K\}. \quad (4.14)$$

ここで, w_T, h_T はそれぞれ T の幅と高さである. また, n_h, n_v はそれぞれ, 原画像における垂直, 水平方向の画像境界の一辺に含まれる格子点の数であり, $\mathbf{1}_{(i)} := (1, 1, \dots, 1)^T \in \mathbb{R}^i$ はすべての要素が1である長さ i のベクトルである. B_{lx} and $B_{rx} \in \mathbb{R}^{2n_v \times 2n}$ は原画像において右端, または左端に位置する格子点の x 座標を返す行列であり, 同様に B_{ty} および $B_{by} \in \mathbb{R}^{2n_h \times 2n}$ は原画像において上端, または下端に位置する格子点の y 座標を返す行列である. また, $\mathbf{u}_i^x$ および $\mathbf{u}_i^y$ は, ある格子点 $\mathbf{u}_i$ のそれぞれ x 座標, y 座標を表す.

前節の H_f と同様の論理から, H_b も凸集合であり, H_l は各要素に対するボックス制約集合となることから, これもまた凸集合である.

4.4 アルゴリズム

提案手法2では、凸最適化問題 (4.5) を解く解法として alternating directions method of multipliers (ADMM) [36, 37] と呼ばれる制約付き凸最適化に対する解法を採用した。本節ではその詳細について述べる。

4.4.1 Alternative Directions Method for Multipliers (ADMM)

いま、以下の最適化問題を考える。

$$\min_{\mathbf{x} \in \mathbb{R}^{N_1}} \phi_1(\mathbf{x}) + \phi_2(M\mathbf{x}), \quad (4.15)$$

ここで、 $\phi_1 \in \Gamma_0(\mathbb{R}^{N_1})$ 、 $\phi_2 \in \Gamma_0(\mathbb{R}^{N_2})$ であり、 $M \in \mathbb{R}^{N_2 \times N_1}$ は任意の行列である。この最適化問題は、新しい変数 $\mathbf{s}(= M\mathbf{x})$ を導入することで、以下のように制約付き最適化問題として定式化できる。

$$\min_{\mathbf{x} \in \mathbb{R}^{N_1}, \mathbf{s} \in \mathbb{R}^{N_2}} \phi_1(\mathbf{x}) + \phi_2(\mathbf{s}), \text{ subject to } \mathbf{s} = M\mathbf{x}. \quad (4.16)$$

この最適化問題を解くために、Alternative Directions Method for Multipliers (ADMM [36, 37, 38]) と呼ばれる反復アルゴリズムが提案されている。そのアルゴリズムを、Algorithm 2 に示す。

また、ADMM の収束性は Eckstein-Bertsekas による以下の定理 4.4.1 によって示される。

定理 4.4.1 (Eckstein-Bertsekas [36]) 式 (4.15) において、 $M^T M$ が正則行列であり、 $\mathcal{L}_0(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \phi_1(\mathbf{x}) + \phi_2(\mathbf{y}) - \langle \mathbf{z}, M\mathbf{x} - \mathbf{y} \rangle$ の鞍点が存在すると仮定する。このとき、Algorithm 2 によって生成される点列 $(\mathbf{x}_i)_{i \in \mathbb{N}}$ は式 (4.15) の解に収束する。

さらに、以下の定理により問題 (4.15) の鞍点の存在性が保証されることを明記しておく。

定理 4.4.2 (Yamagishi-Ono-Yamada [39]) 式 (4.15) において、この問題が少なくとも一つの解をもち、行列 $M^T M$ が正則であり、 $0 \in \text{ri}(\text{dom}(\phi_2) - M\text{dom}(\phi_1))$ と仮定する。このとき $\mathcal{L}_0$ の鞍点が存在する。ここで $\text{ri}(\cdot)$ は $(\cdot)$ の相対的内部を表す。

Algorithm 2 ADMM [36, 37]

1: $m = 0$, and choose $\mu > 0$, $\mathbf{s}_0$, $\mathbf{t}_0$.
2: **repeat**3: $\mathbf{x}_{m+1} \in \arg \min_{\mathbf{x} \in \mathbb{R}^{N_1}} \phi_1(\mathbf{x}) + (\mu/2) \|\mathbf{M}\mathbf{x} - \mathbf{s}_m - \mathbf{t}_m\|_2^2.$ 4: $\mathbf{s}_{m+1} \in \arg \min_{\mathbf{s} \in \mathbb{R}^{N_2}} \phi_2(\mathbf{s}) + (\mu/2) \|\mathbf{M}\mathbf{x}_{m+1} - \mathbf{s} - \mathbf{t}_m\|_2^2.$ 5: $\mathbf{t}_{m+1} = \mathbf{t}_m - (\mathbf{M}\mathbf{x}_{m+1} - \mathbf{s}_{m+1}).$ 6: $m \leftarrow m + 1$ 7: **until** some stopping criterion is satisfied.

4.4.2 ADMMの拡張

既存研究 [38] により ADMM は、以下のようなより一般的な問題にも適用可能であることが示されている。

$$\min_{\mathbf{x} \in \mathbb{R}^{N_1}} \psi(\mathbf{x}) + \sum_{j=1}^J \tilde{\psi}^{(j)}(\mathbf{M}^{(j)}\mathbf{x}), \quad (4.17)$$

ここで $\psi \in \Gamma_0(\mathbb{R}^{N_1})$ は微分可能な関数であり、 $\tilde{\psi}_j \in \Gamma_0(\mathbb{R}^{N_2^{(j)}})$ はそれらの近接作用素が閉形式解として与えられる関数である。さらに、 $\mathbf{M}^{(j)} \in \mathbb{R}^{N_2^{(j)} \times N_1}$ は任意の行列である。この最適化問題は式 (4.16) と以下のように対応している。

$$\phi_1(\mathbf{x}) = \psi(\mathbf{x}), \quad (4.18)$$

$$\phi_2(\mathbf{s}) = \sum_{j=1}^J \tilde{\psi}^{(j)}(\mathbf{s}^{(j)}), \quad (4.19)$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^{(1)} \\ \vdots \\ \mathbf{M}^{(J)} \end{bmatrix} \in \mathbb{R}^{N_2 \times N_1}, \quad (4.20)$$

ここで、 $N_2 = N_2^{(1)} + \dots + N_2^{(J)}$ 、 $\mathbf{s}^{(j)} \in \mathbb{R}^{N_2^{(j)}}$ 、 $\mathbf{s} = [(\mathbf{s}^{(1)})^T, (\mathbf{s}^{(2)})^T, \dots, (\mathbf{s}^{(J)})^T]^T$ である。この対応関係により、ADMM は容易に式 (4.17) に適用可能である。まず、

Algorithm 2 の Step 3 は以下のように定式化可能である.

$$\begin{aligned}
\mathbf{x}_{m+1} &\in \arg \min_{\mathbf{x} \in \mathbb{R}^{N_1}} \psi(\mathbf{x}) + (\mu/2) \|\mathbf{M}\mathbf{x} - \mathbf{s}_m - \mathbf{t}_m\|_2^2 \\
&= \arg \min_{\mathbf{x} \in \mathbb{R}^{N_1}} \psi(\mathbf{x}) \\
&\quad + \frac{\mu}{2} \left\| \begin{bmatrix} M^{(1)} \\ \vdots \\ M^{(J)} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{s}_m^{(1)} \\ \vdots \\ \mathbf{s}_m^{(J)} \end{bmatrix} - \begin{bmatrix} \mathbf{t}_m^{(1)} \\ \vdots \\ \mathbf{t}_m^{(J)} \end{bmatrix} \right\|_2^2 \\
&= \arg \min_{\mathbf{x} \in \mathbb{R}^{N_1}} \psi(\mathbf{x}) + (\mu/2) \sum_{j=1}^J \|M^{(j)}\mathbf{x} - \mathbf{s}_m^{(j)} - \mathbf{t}_m^{(j)}\|_2^2. \tag{4.21}
\end{aligned}$$

ψ は微分可能であるため, 式 (4.21) の解は以下の式を解くことで容易に得られる.

$$\frac{\partial}{\partial \mathbf{x}} \left[\psi(\mathbf{x}) + (\mu/2) \sum_{j=1}^J \|M^{(j)}\mathbf{x} - \mathbf{s}_m^{(j)} - \mathbf{t}_m^{(j)}\|_2^2 \right] = 0. \tag{4.22}$$

一方, Algorithm 2 の Step 4 は以下のように書くことができる.

$$\begin{aligned}
\begin{bmatrix} \mathbf{s}_{m+1}^{(1)} \\ \vdots \\ \mathbf{s}_{m+1}^{(J)} \end{bmatrix} &\leftarrow \arg \min_{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(J)}} \tilde{\psi}^{(1)}(\mathbf{s}^{(1)}) + \dots + \tilde{\psi}^{(J)}(\mathbf{s}^{(J)}) \\
&\quad + \frac{\mu}{2} \left\| \begin{bmatrix} M^{(1)} \\ \vdots \\ M^{(J)} \end{bmatrix} \mathbf{x}_{m+1} - \begin{bmatrix} \mathbf{s}^{(1)} \\ \vdots \\ \mathbf{s}^{(J)} \end{bmatrix} - \begin{bmatrix} \mathbf{t}_m^{(1)} \\ \vdots \\ \mathbf{t}_m^{(J)} \end{bmatrix} \right\|_2^2 \\
&= \arg \min_{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(J)}} \tilde{\psi}^{(1)}(\mathbf{s}^{(1)}) + \dots + \tilde{\psi}^{(J)}(\mathbf{s}^{(J)}) \\
&\quad + (\mu/2) \sum_{j=1}^J \|M^{(j)}\mathbf{x}_{m+1} - \mathbf{s}^{(j)} - \mathbf{t}_m^{(j)}\|_2^2. \tag{4.23}
\end{aligned}$$

さらに, 以下のそれぞれの式を独立に解くことで, $\mathbf{s}_{m+1}^{(j)}$ ($j \in \{1, \dots, J\}$) が得られる.

$$\mathbf{s}_{m+1}^{(j)} = \arg \min_{\mathbf{s}^{(j)} \in \mathbb{R}^{N_2^{(j)}}} \tilde{\psi}^{(j)}(\mathbf{s}^{(j)}) + (\mu/2) \|M^{(j)}\mathbf{x}_{m+1} - \mathbf{s}^{(j)} - \mathbf{t}_m^{(j)}\|_2^2. \tag{4.24}$$

この式は近接作用素を用いて以下のように書き直すことができる.

$$\mathbf{s}_{m+1}^{(j)} = \text{prox}_{\tilde{\psi}^{(j)}/\mu}(M^{(j)}\mathbf{x}_{m+1} - \mathbf{t}_m^{(j)}). \tag{4.25}$$

Algorithm 3 ADMM-2 [36, 37, 38]

-
- 1: $m = 0$, and choose $\mu > 0$, $s_0^{(1)}, \dots, s_0^{(J)}, t_0^{(1)}, \dots, t_0^{(J)}$.
 - 2: **repeat**
 - 3: $\mathbf{x}_{m+1}^{(j)} \in \arg \min_{\mathbf{x} \in \mathbb{R}^{N_1}} \psi(\mathbf{x}) + (\mu/2) \sum_{j=1}^J \|M^{(j)}\mathbf{x} - \mathbf{s}_m^{(j)} - \mathbf{t}_m^{(j)}\|_2^2$
 (solve Eq. (4.22)).
 - 4: **for** $i = 1, \dots, J$ **do**
 - 5: $\mathbf{s}_{m+1}^{(j)} \in \text{prox}_{\tilde{\psi}^{(j)}/\mu}(M^{(j)}\mathbf{x}_{m+1} - \mathbf{t}_m^{(j)})$.
 - 6: $\mathbf{t}_{m+1}^{(j)} = \mathbf{t}_m^{(j)} - (M^{(j)}\mathbf{x}_{m+1} - \mathbf{s}_{m+1}^{(j)})$.
 - 7: **end for**
 - 8: $m \leftarrow m + 1$
 - 9: **until** some stopping criterion is satisfied.
-

式 (4.21) と式 (4.25) から, 式 (4.17) のための ADMM は Algorithm 3 (ADMM-2) のようになる.

4.4.3 提案手法 2 における ADMM

前節で述べた ADMM を提案手法 2 のエネルギー汎関数 (4.5) に適用する. 式 (4.5) と 式 (4.18)~(4.20) における各要素の対応関係は以下のとおりとなる.

$$\psi(\mathbf{u}) = \frac{\lambda_2}{2} \|\mathbf{Gu} - \mathbf{Gv}\|_2^2 + \frac{\lambda_3}{2} \|\mathbf{Du}\|_2^2, \quad (4.26)$$

$$\begin{aligned} \tilde{\psi}^{(1)}(\mathbf{s}^{(1)}) &= \frac{\lambda_1}{2} d_{A_i}^2(\mathbf{s}^{(1)}), \quad \tilde{\psi}^{(2)}(\mathbf{s}^{(2)}) = \iota_{W_i}(\mathbf{s}^{(2)}), \\ \tilde{\psi}^{(3)}(\mathbf{s}^{(3)}) &= \iota_{H_f}(\mathbf{s}^{(3)}), \quad \tilde{\psi}^{(4)}(\mathbf{s}^{(4)}) = \iota_{H_b}(\mathbf{s}^{(4)}), \\ \tilde{\psi}^{(5)}(\mathbf{s}^{(5)}) &= \iota_{H_l}(\mathbf{s}^{(5)}), \end{aligned} \quad (4.27)$$

$$\mathbf{M} = \left[\mathbf{G}^T \mathbf{D}^T \mathbf{R}_1^T \cdots \mathbf{R}_k^T \mathbf{R}_1^T \cdots \mathbf{R}_k^T \mathbf{F}^T \mathbf{B}^T \mathbf{U}^T \right]^T, \quad (4.28)$$

ここで $\mathbf{U}$ は $2n \times 2n$ の単位行列であり, W_i はすべてのインデックス j に対する $W_{i,j}$ の積集合である ($W_i := \bigcap_{j \in K \setminus i} W_{i,j}$).

提案手法 2 における ADMM の詳細を Algorithm 4 に示す. いま, H_f , H_b , H_l , および W_i は凸集合となっているため, その指示関数 ι_{H_f} , ι_{H_b} , ι_{H_l} , および ι_{W_i} の近接作用素は, 各集合に対する距離射影を取ることと等しく, これらは容易に計算可能である. また, ある集合 C に対する距離関数 $(\lambda/2\mu)d_C^2$ の近接作用

素は以下のように計算される.

$$\text{prox}_{(\lambda/2\mu)d_C^2}(\mathbf{x}) = \frac{(\lambda/\mu)P_C(\mathbf{x}) + \mathbf{x}}{(\lambda/\mu) + 1}. \quad (4.29)$$

これらの条件により, 式 (4.5) の最小解の存在性と Algorithm 4 の収束性に関する以下の定理が導出される.

定理 4.4.3 (式 (4.5) の最小解の存在性) 式 (4.5) における最小解 $\mathbf{u}^* \in \mathbb{R}^{2n}$ が存在する.

証明の詳細は付録 I.1 に示す.

定理 4.4.4 (Algorithm 4 の収束性) Algorithm 4 によって導出される点列 $(\mathbf{u}_m)_{m \in \mathbb{N}}$ は式 (4.5) の最小解である $\mathbf{u}^*$ に収束する.

証明の詳細は付録 I.2 に示す.

4.5 実験

4.5.1 実験環境

本章では, 複数の画像に対する提案手法2と既存手法との比較実験を行う. 図 4.5~4.7² に実験結果を示す. これらはすべて, 原画像より幅の狭い任意形ディスプレイで画像を表示した際の結果である. 比較対象とする既存手法としては, “Seam Carving (SC) [1]”, “Image Summarization via Bidirectional Similarity (ISBS) [2]”, “提案手法2:ピクセルベースの任意形イメージリターゲッティング (Pixel based Arbitrary-shaped Image Retargeting, PAIR) ”, “Axis-Aligned Deformation (AAD) [7]”, “Feature-Aware Texturing (FAT) [3]” の5つの手法を採用した. SC, ISBS, PAIR, AAD, FAT はすべて2章において述べた既存のイメージリターゲッティング手法であり, SC, ISBS, PAIR はピクセルベース, AAD と FAT はワーピングベースのアプローチに分類される.

既存手法についてはすべて, 提案手法2と同様に, 任意形の表示領域に外接するサイズの矩形の画像を出力し, それらに対して表示領域の形状を表す

²The authors are permitted to publish and modify of the photographs released under the Creative Commons. The photograph named “Deer” is from Noël Zia Lee, “Cars” is from Noel C. Hankamer, and “Rope skipping” is from Jose Maria Cuellar.

Algorithm 4 Specialized ADMM for solving Eq. (4.5)

-
- 1: Set $m = 0$, and choose $\mu > 0$, $\mathbf{s}_0^{(f)}$, $\mathbf{t}_0^{(f)}$, $\mathbf{s}_0^{(b)}$, $\mathbf{t}_0^{(b)}$, $\mathbf{s}_0^{(l)}$, $\mathbf{t}_0^{(l)}$, $\mathbf{s}_{i,0}^{(a)}$, $\mathbf{t}_{i,0}^{(a)}$, $\mathbf{s}_{i,0}^{(w)}$, $\mathbf{t}_{i,0}^{(w)}$ ($i \in K$).
 - 2: $M = \lambda_2 G^T G + \lambda_3 D^T D + \mu F^T F + \mu B^T B + 2\mu \sum_{i \in K} R_i^T R_i + \mu U$.
 - 3: **repeat**
 - 4: $\mathbf{r}_m = \lambda_2 G^T G \mathbf{v} + \mu F^T (\mathbf{s}_m^{(f)} + \mathbf{t}_m^{(f)}) + \mu B^T (\mathbf{s}_m^{(b)} + \mathbf{t}_m^{(b)}) + \mu (\mathbf{s}_m^{(l)} + \mathbf{t}_m^{(l)}) + \mu \sum_{i \in K} R_i^T (\mathbf{s}_{i,m}^{(a)} + \mathbf{t}_{i,m}^{(a)}) + \mu \sum_{i \in K} R_i^T (\mathbf{s}_{i,m}^{(w)} + \mathbf{t}_{i,m}^{(w)})$.
 - 5: $\mathbf{u}_{m+1} = M^{-1} \mathbf{r}_m$.
 - 6: $\mathbf{s}_{m+1}^{(f)} = \text{prox}_{\iota_{H_f}} (F \mathbf{u}_{m+1} - \mathbf{t}_m^{(f)})$.
 - 7: $\mathbf{t}_{m+1}^{(f)} = \mathbf{t}_k^{(f)} - F \mathbf{u}_{m+1} + \mathbf{s}_{m+1}^{(f)}$.
 - 8: $\mathbf{s}_{m+1}^{(b)} = \text{prox}_{\iota_{H_b}} (B \mathbf{u}_{m+1} - \mathbf{t}_m^{(b)})$.
 - 9: $\mathbf{t}_{m+1}^{(b)} = \mathbf{t}_k^{(b)} - B \mathbf{u}_{m+1} + \mathbf{s}_{m+1}^{(b)}$.
 - 10: $\mathbf{s}_{m+1}^{(l)} = \text{prox}_{\iota_{H_l}} (\mathbf{u}_{m+1} - \mathbf{t}_m^{(l)})$.
 - 11: $\mathbf{t}_{m+1}^{(l)} = \mathbf{t}_k^{(l)} - \mathbf{u}_{m+1} + \mathbf{s}_{m+1}^{(l)}$.
 - 12: **for all** i such that $i \in K$ **do**
 - 13: $\mathbf{s}_{i,m+1}^{(a)} = \text{prox}_{(\lambda_1/2\mu)d_{\Lambda_i}^2} (R_i \mathbf{u}_{m+1} - \mathbf{t}_{i,m}^{(a)})$.
 - 14: $\mathbf{t}_{i,m+1}^{(a)} = \mathbf{t}_{i,m}^{(a)} - R_i \mathbf{u}_{m+1} + \mathbf{s}_{i,m+1}^{(a)}$.
 - 15: $\mathbf{s}_{i,m+1}^{(w)} = \text{prox}_{\iota_{W_i}} (R_i \mathbf{u}_{m+1} - \mathbf{t}_{i,m}^{(w)})$.
 - 16: $\mathbf{t}_{i,m+1}^{(w)} = \mathbf{t}_{i,m}^{(w)} - R_i \mathbf{u}_{m+1} + \mathbf{s}_{i,m+1}^{(w)}$.
 - 17: **end for**
 - 18: $m \leftarrow m + 1$
 - 19: **until** some stopping criterion is satisfied.
-

マスクをかけた結果となっている（図 4.7 については矩形の表示領域であるため、マスク処理を行わないそのままの画像出力である）。特に、FAT に関しては、原画像の画素値をすべて任意形の表示領域内部にマッピングするように画像を出力することも可能であるが、図 2.6 からわかるとおり、画像境界に近い領域が大きく歪むことがわかっている。そこで、本提案手法 2 との公平性を保つため、FAT においても、任意形の表示領域に外接する矩形領域の画像出力を行ったのち、表示領域の形状に基づくマスク処理を行っている（FAT は矩形画像出力もサポートしており、本実験において特別な処理の変更は行っていないことに注意されたい）。

また、ISBS を除く既存手法（SC, PAIR, AAD, FAT）では、その処理のために原画像内の画素の重要度を決定する必要がある。これについては本手法

同様，Hou らの顕著性マップ [26] を採用した．さらに，これらの中でも PAIR と FAT については，画像内の重要な領域とそうでない領域を分割する 2 値化処理を行う必要があるが，この処理についても本手法同様，閾値 κ を用いた 2 値化処理を採用した（提案手法 2 では 2 値化ののちに画素集合の凸包を計算するが，既存手法に対しては，2 値化処理のみを行っている）．その際の各画像に対する閾値は表 4.1 のとおりであり，これらの閾値は発見的に決定した．同一の画像に対する閾値 κ は手法によらず一定であることに注意されたい．

表 4.1 各画像における閾値 κ

入力画像の名称	κ
Fish (heart)	0.5
Cars (heart)	0.5
Deer (heart)	0.5
Dolphins (heart)	0.4
Fish (diamond)	0.5
Cars (spade)	0.4
Dolphins (club)	0.4
Deer (rect.)	0.5
Rope skipping (rect.)	0.3
Cliff (rect.)	0.6

また，提案手法 2 のエネルギー汎関数における各項の重み付け係数は，すべての画像において $\lambda_1 = 0.01$ ， $\lambda_2 = 0.1$ ， $\lambda_3 = 1$ であり， κ と同様に発見的な決定によるものであるが，同桁数内でのパラメータの変化は結果に大きく影響を与えないことを確認している．格子間隔は水平，垂直方向ともに 10 [pixel] である．

4.5.2 考察

図 4.5～4.7 の結果から，提案手法 2 では，表示領域の形状に合わせた ROI の移動とその形状の保存が行われており，ROI がより大きく表示されていることがわかる．これと比較して既存手法の結果に対する考察を行う．

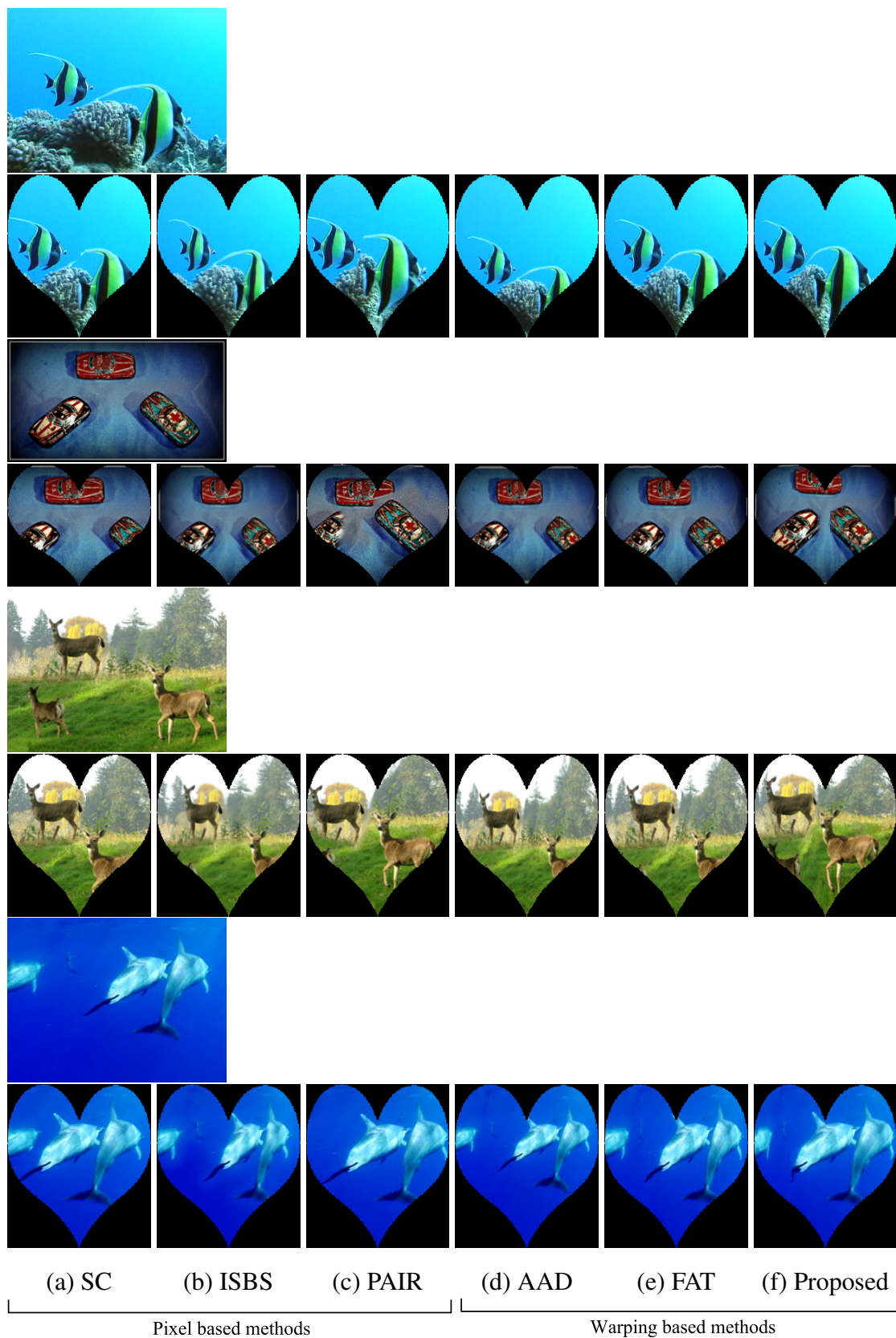


図 4.5 ハート形の表示領域における既存手法との比較（画像の名称は上から下へ順に，“Fish (heart)”，“Cars (heart)”，“Deer (heart)”，“Child (heart)”であり，各行の左上の画像が原画像である。）

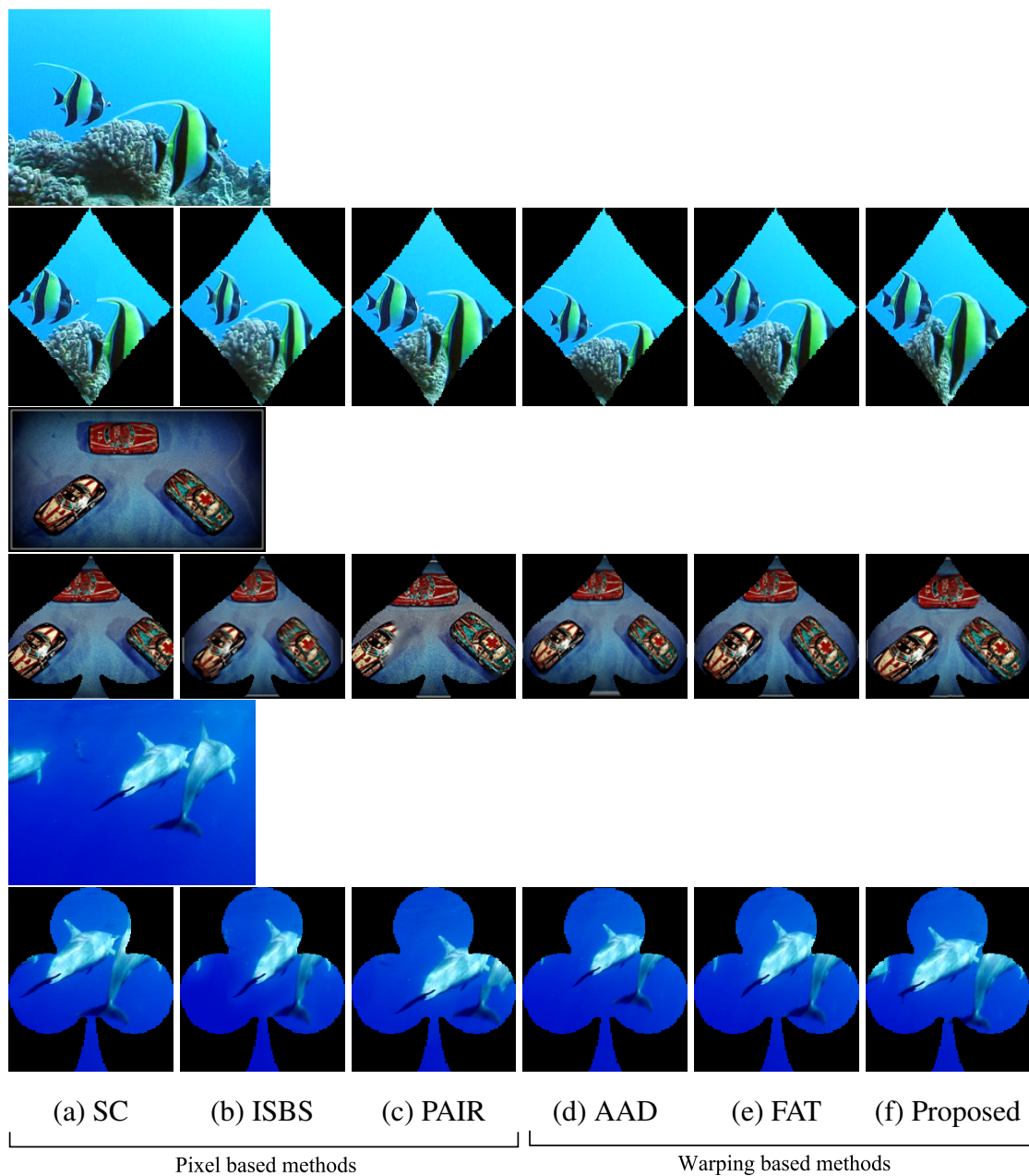


図 4.6 多様な形状の表示領域における既存手法との比較（画像の名称は上から下へ順に，“Fish (diamond)”，“Cars (spade)”，“Dolphins (club)”であり，各行の左上の画像が原画像である。）

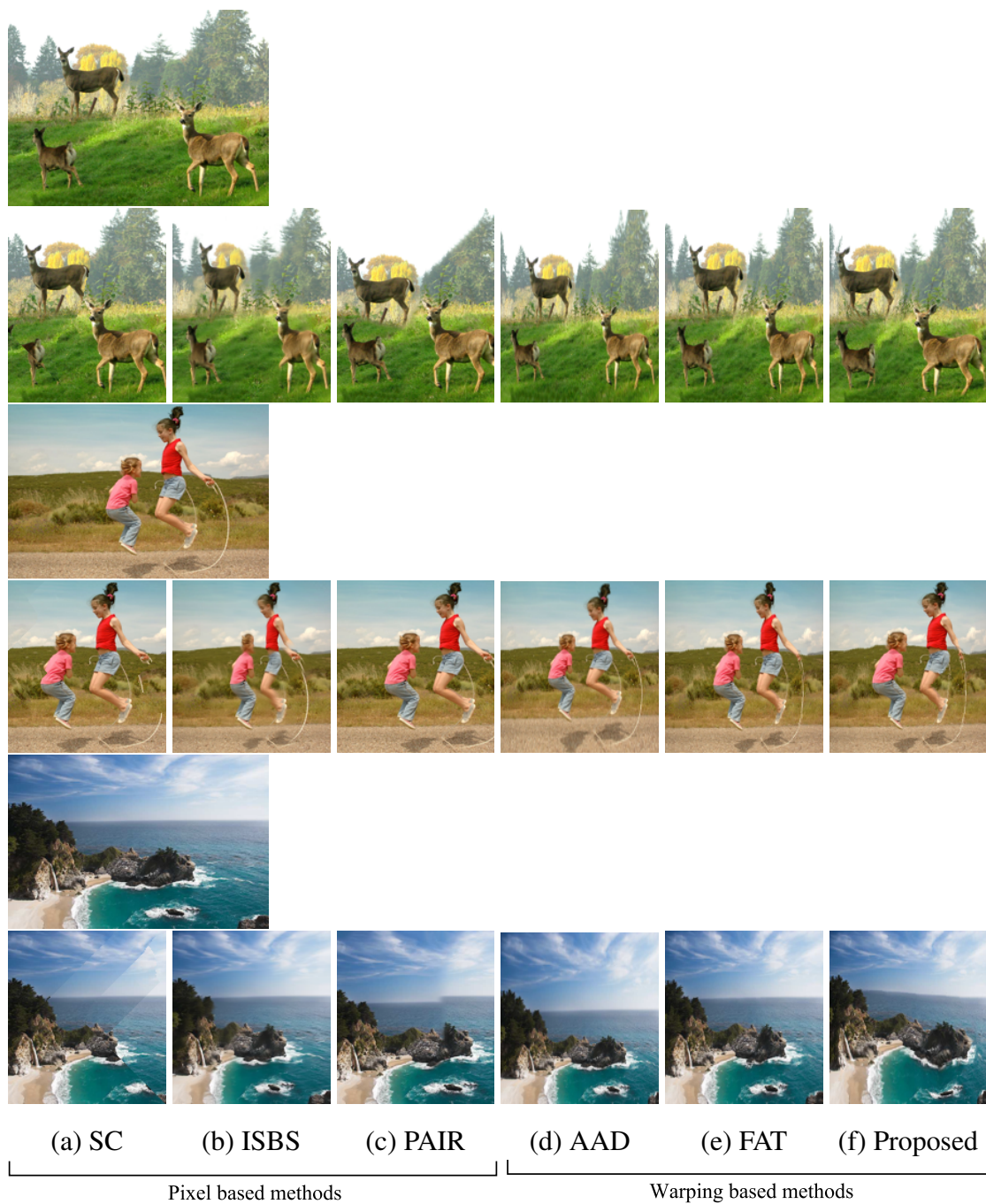


図 4.7 従来までの対象である矩形の表示領域における既存手法との比較（画像の名称は上から下へ順に，“Deer (rect.)”，“Rope skipping (rect.)”，“Cliff (rect.)”であり，各行の左上の画像が原画像である。）

まず SC では、不要な画素を削除する過程で ROI 周辺の画素も削除してしまう可能性がある。このため、図 4.5, 4.6 の Fish では、右の魚のヒレが途中で途切れてしまっていることがわかる。これと同様の現象は図 4.7 の Rope skipping のロープにも確認できる。また、図 4.5 の Cars では左右の車の大半が、Deer では左下のシカが表示されていないといったように、画像内の ROI を表示領域内で表示できていないケースが多くみられる。

続いて、ISBS の結果画像では、画像全体のアスペクト比の変化にともない ROI 自体のアスペクト比も変化していることがわかる。例えば図 4.5, 4.6 の Cars では上部の車の長さが極端に変化している。また、図 4.7 の Rope skipping において大きく人物が歪んでしまっている。これらは、不自然な印象を与えるため、望ましい結果であるとはいえない。また、SC 同様に画像内の ROI を表示しきれていない結果が多い。

PAIR は、任意形の表示領域を対象としたイメージリターゲティングであり、ROI の移動を可能とする手法である。結果画像においても、図 4.5 の Fish では、ROI が大きく移動していることがわかる。しかしながら、図 4.5, 4.6 の Dolphins では左端のイルカが全く表示されていない。この問題は、PAIR が採用している発見的アルゴリズムが局所最適解を導出した結果、イルカの移動が、より望ましい位置の手前で止まってしまったことに起因すると考えられる。また、図 4.5, 4.6 の Cars では車の境界に歪み（不自然なぼけ）が生じている。PAIR は ISBS と同様に原画像に含まれるブロックの線形和で画像を表現する手法であり、これらの手法では本質的にブロックの重ね合わせに起因するぼけを回避することは難しい。

AAD, FAT は提案手法 2 と同様のワーピングベースの手法であり、すべての結果画像において大きく目立つ断裂やぼけは起こっていない。しかし、ROI の移動を可能としないため、図 4.5 の fish における右の魚や、Deer における左のシカなど、提案手法 2 に比べて ROI が適切に表示されていない結果が多い。

以上の考察から、提案手法 2 は既存手法に比べ、任意形の表示領域の形状に合わせた ROI の移動と、その ROI の形状の保存を可能としたイメージリターゲティングが実現されているといえる。

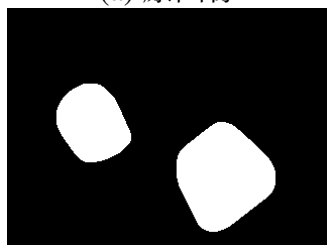
今回の実験では、ROI を決定する際の閾値 κ は発見的に決定しており、その最適性は保証されていない。誤った閾値設定がなされた際に問題となるのは、前景オブジェクトなどの本来 ROI とされるべき画像内の領域が適切に ROI とされない場合である（PAIR, FAT にも同様の問題が起こりうることに注意さ

りたい)。しかし、提案手法2ではワーピングベースのアプローチを採用しており、この場合でも格子点同士の連結性は失われないため、オブジェクト自体が完全に失われてしまうといったことは考えにくい（図4.8）。また、同様の理由からSCやPAIRなどに比べて、明らかな断裂は起こしにくく、結果として歪みの少ない画像が得られると考えられる。しかしながら、図4.8からもわかるとおり、この例では提案手法2の意図するROIの移動の効果は損なわれている。

この問題を解決する方策として、最適な閾値設定手法の開発や、領域分割を組み合わせたROI抽出手法、ユーザー入力を手がかりとするROI抽出手法などが考えられる。これらは今後の課題であるといえる。



(a) 原画像



(b) ROI ($\kappa = 0.5$)



(c) 結果画像
($\kappa = 0.5$)



(d) ROI ($\kappa = 0.7$)



(e) 結果画像
($\kappa = 0.7$)

図 4.8 閾値 κ による結果の違い（図 4.5 の結果は (c)）

4.6 むすび

本章では、提案手法2として、ワーピングベースの任意形イメージリタージェットィングを提案した。

任意形の表示領域を対象とした際の主要な課題は、ROIをどの位置に配置するべきか、という問題を解くことである。提案手法2では、ワーピングベー

スのアプローチを応用し、ROIの移動問題を、原画像を格子状に分割した際の格子点座標を決定する問題として扱っている。その上で、ROIの移動可能範囲をあらかじめ制限することで、格子点移動のためのエネルギー汎関数を凸関数として定式化し、大域最適解としての格子点の移動を可能としている。このROI移動の大域最適性は、従来の任意形イメージリターゲットングにはない特徴であり、移動が局所最適解に陥ることによりROIが表示されないといった問題を防ぐことができる。

実験では、提案手法2がROIの移動と、その形状を保存した画像表示を可能とし、任意形の表示領域において有効であることが確認された。それに対して、既存手法を任意形の表示領域に対して適用した場合には、ROIが表示領域内に収まらず、表示できない場合があるという結果が得られた。また、一部の既存手法においては、ROIの不自然な断裂やぼけが生じることが確認された。

本提案手法2の性能は、ROI抽出の結果の影響を大きく受けるため、誤ったROIが抽出された場合は、結果画像中に不自然な歪みが生じる可能性がある。しかしながら、このような場合でも、ワーピングベースのアプローチを応用したことにより、提案手法2ではオブジェクトの欠損などが起こりにくいことも示した。

第5章

結論

本論文では、任意形の表示領域をもつデバイスでの画像表示の際に起こる問題に対して検討を行い、それを解決する方法として、任意形のイメージリターゲッティングを提案した。

まず、矩形の表示領域を対象としてきた既存のイメージリターゲッティングが扱ってきた、「ROIの形状を保存しながら表示領域に収まるように移動」させつつ、同時に「画像全体の歪みを抑え違和感のない画像生成を行う」といった問題について検討を行い、それを任意形の表示領域に拡張する際の問題点について述べた。そして、既存手法が、そのアプローチによりピクセルベースとワーピングベースという2つの分類に大別できることを提示し、それらの特徴についてまとめた上で、それぞれに対応する2つの手法を提案した。

第3章では、提案手法1として、既存のイメージリターゲッティング手法[2]において提案された客観評価指標Bidirectional Similarityを考慮した、ピクセルベースのイメージリターゲッティング手法を提案した。この手法では、処理のプロセスを2段階に分割することにより、任意形の表示領域に合わせた画像表示を可能としている。また、各プロセスにおいてBidirectional Similarityを考慮することで、客観評価指標の上でも良好な手法を実現している。しかしながら、ピクセルベースの手法の特徴である、最適化問題が非凸となってしまう出力される解が局所最適に陥ってしまうという問題は、提案手法1にも同様に存在する。

そこで第4章では、提案手法2として、ROI移動の大域最適化によるワーピングベースの任意形イメージリターゲッティングを提案した。ワーピングベースのイメージリターゲッティングは、原画像を一定間隔の格子状に分割した



(a) 原画像

(b) 分割された任意形の領域それぞれにリターゲットイングによる拡大表示を行った結果

図 5.1 任意形イメージリターゲットイングのアプリケーション応用例

際に得られる格子点の、座標の決定問題として扱うことができる。提案手法2では、このワーピングベースの枠組みに基づき、その上で、ROI 移動に対して原画像から算出される移動可能範囲を設定することで、凸なエネルギー汎関数の設計を可能とした。このことにより、定められたエネルギー汎関数において、大域最適となる任意形イメージリターゲットイングを実現した。

以上、本論文で得られた成果をまとめると次のようになる。任意形の表示領域という形状の制約のない一般化された問題を扱うことで、イメージリターゲットイングが内包する問題の所在が何であるのかを明らかにした。その上で、既存手法の分類にそれぞれ対応する2つの手法を提案することにより、これまでのイメージリターゲットイングに対する包括的な問題解決を行った。これにより、将来的に応用が期待される様々な形状をもつディスプレイデバイスにおいて、矩形の画像表示を行うことが可能となった。

さらに、本研究の成果は任意形のデバイスだけに限らず、新しい画像アプリケーションへの応用を可能とする。例えば図 5.1 のように、矩形の画像を、その内容に応じていくつかの任意形の領域に分割し、各領域ごとにイメージリターゲットイングによる拡大表示を行うといったことが可能となる。このようにデバイス、アプリケーション共に、表示領域の形状の制約を一般化することで、新たな画像表示が期待できる。これは、今後の画像応用の可能性を広げることに繋がる。

本論文において提案された手法の将来的な展望として、実用化を目指した動画への適用が挙げられる。この目的の実現のためには主に2点の課題が挙

げられる．時刻変化に対する自然な動きの実現と処理時間の高速化である．

まず，単純に動画のフレームごとに本提案手法を適用しただけでは，時間方向の連結性が考慮されていないためオブジェクトの動きは不自然なものになってしまう．そのため，自然な動きを実現するために，本論文で提案されたエネルギー汎関数にさらに時間方向の拘束条件を導入することが考えられる．これは直前のフレームの結果を初期値として利用するなどの工夫により，より簡易に設計可能となることが予想される．また，この直前のフレームの結果の利用などは，フレームごとの処理時間を削減することにもつながり，全体として大幅な処理時間の高速化が可能となることも予想される．このように，今後の研究課題として本論文の動画への適用を行うことは十分可能であると考えられる．

また，ユーザーが真に望む画像表示は，画像の使用される状況によって様々である．提案手法も含めたこれまでのイメージリターゲッティングは，ユーザー入力による ROI 決定により，より一層望ましい画像を得ることを可能としている．そして今後は，このようなユーザーとのインタラクティブをまったく必要としないイメージリターゲッティングが必要とされることが予想される．そのために，任意形イメージリターゲッティングにおいても，用途に応じたパラメーター設定法，ROI 決定法などが必要であり，これは今後の研究課題であるといえる．同時に，実際のアプリケーション応用を提案し，ユーザーによる評価を行うことも今後の研究課題である．

謝辞

本研究を進めるにあたり、終始一貫して丁寧なご指導ご鞭撻を賜りました東京工業大学 羽鳥好律教授に感謝の意を申し上げます。博士課程からの研究室所属ということで、不安な面も多々あったかと思いますが、温かく迎え入れ、常に見守ってくださいましたことを心より感謝しております。羽鳥先生のもとで、研究に対する考え方や姿勢、専門知識など、多くのご教授を賜りましたことは、今後の人生においても大きな糧となるものであります。

修士課程時の私の指導教員であらせられます放送大学 酒井善則特任教授（当時、東京工業大学教授）には、博士課程進学にあたって多大なご支援を賜りました。研究に対するご指導だけでなく、人生の岐路において温かいご助言を賜りましたことは、一生の財産となるものであります。

つぎに、学位論文審査の労をお執りくださいました東京工業大学 長橋宏教授、小林隆夫教授、小池康晴教授、佐藤いまり連携准教授、千葉工業大学工学部 宮田高道准教授に感謝申し上げます。宮田高道先生には、修士課程時から長年に渡り、非常に熱心なご指導を賜りました。論文執筆にあたってのご指導、技術および理論に関するご助言など、言葉には尽くせないほどのご助力を賜り、併せて深く感謝申し上げます。

東京工業大学大学院 集積システム専攻 博士課程学生の小野峻佑さんには、たびたび貴重なご討論をさせていただき、特に理論に関する有益なアドバイスをいただきました。また、同じ博士課程学生ということで、常に新しい刺激をいただき、研究活動を遂行する上での原動力となりました。感謝申し上げます。

さらに、研究生を送る上での様々なご支援、研究上の有益なご意見をいただきました羽鳥研究室の皆様へ感謝致します。多くの良き先輩、後輩の方々に恵まれ、交流する中で得られた経験は、貴重な財産を私に与えてくれました。また、羽鳥研究室 秘書 安藤眞弓様には、事務手続きにあたって多くのご助力を賜りました。皆様のおかげで、研究に専念することができ、また、

楽しい研究生生活を送ることができました。

文部科学省「グローバルCOEプログラム“フォトニクス集積コアエレクトロニクス”」および「卓越した大学院拠点形成支援補助金プログラム」、東京工業大学 総合理工学研究科リサーチアシスタント経費により、研究生生活を送る上での金銭的なご支援を賜りました。感謝申し上げます。

3年に渡る博士課程学生生活において、多くの楽しく心地よい時間を過ごし、日々の活力を与えてくれた友人の、渡邊和也さん、橘内健浩さん、高岡敦さん、貝原亮太さん、氏野彰信さん、斎藤隆さんに感謝申し上げます。

最後に、学位取得にあたり常に見守り、精神的な支えとなってくれた両親、姉、兄に感謝の意を表し、本論文を捧げます。

参考文献

- [1] S. Avidan and A. Shamir, “Seam carving for content-aware image resizing,” *ACM Trans. Graphics*, vol.26, no.3, 2007.
- [2] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, “Summarizing visual data using bidirectional similarity,” *Proc. CVPR, IEEE Conference on*, pp.1–8, 2008.
- [3] R. Gal, O. Sorkine, and D. Cohen-Or, “Feature-aware texturing,” *Proc. the 17th Eurographics conference on Rendering Techniques*, pp.297–303, 2006.
- [4] M. Rubinstein, A. Shamir, and S. Avidan, “Improved seam carving for video re-targeting,” *ACM Trans. Graphics*, vol.27, no.3, 2008.
- [5] A. Shamir and S. Avidan, “Seam carving for media retargeting,” *Commun. ACM*, vol.52, no.1, pp.77–85, 2009.
- [6] Y. Pritch, E. Kav-Venaki, and S. Peleg, “Shift-map image editing,” *Proc. ICCV, IEEE Conference on*, pp.151–158, 2009.
- [7] D. Panozzo, O. Weber, and O. Sorkine, “Robust image retargeting via axis-aligned deformation,” *Computer Graphics Forum*, vol.31, no.2, pp.229–236, 2012.
- [8] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, “Optimized scale-and-stretch for image resizing,” *ACM Trans. Graphics*, vol.27, no.5, 2008.
- [9] D. Wang, G. Li, W. Jia, and X. Luo, “Saliency-driven scaling optimization for image retargeting,” *The Visual Computer*, vol.27, no.9, pp.853–860, 2011.
- [10] D. Pavic and L. Kobbelt, “Two-colored pixels,” *Computer Graphics Forum*, vol.29, no.2, pp.743–752, 2010.
- [11] Z. Karni, D. Freedman, and C. Gotsman, “Energy-based image deformation,” *Proc. the Symposium on Geometry Processing*, pp.1257–1268, 2009.

- [12] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, “Image retargeting using mesh parametrization,” *Multimedia, IEEE Trans.*, vol.11, no.5, pp.856–867, 2009.
- [13] Y. Jin, L. Liu, and Q. Wu, “Nonhomogeneous scaling optimization for realtime image resizing,” *The Visual Computer*, vol.26, no.6-8, pp.769–778, 2010.
- [14] L. Wolf, M. Guttman, and D. Cohen-Or, “Non-homogeneous content-driven video-retargeting,” *Proc. ICCV. IEEE Conference on*, pp.1–6, 2007.
- [15] “Google glass”. <http://www.google.com/glass/start/>
- [16] S.J. Battersby, “Non rectangular display device,” US Patent 7253865 B2 , 2007.
- [17] S. Morita, “Liquid crystal display device,” US Patent 20100141874 A1 , 2010.
- [18] K. Takatori, “Non-rectangular pixel array and display device having same,” US Patent 20120176051 A1 , 2012.
- [19] 高取憲一, 野中義弘, “任意形状の液晶ディスプレイ,” *映情学誌*, vol.64, no.10, pp.1437–1441, 2010.
- [20] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, “Fast image recovery using variable splitting and constrained optimization,” *Trans. Img. Proc.*, vol.19, no.9, pp.2345–2356, 2010.
- [21] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *PAMI, IEEE Trans.*, vol.20, no.11, pp.1254–1259, 1998.
- [22] L. Zhang, M. Tong, H. Marks, T. and Shan, and G. Cottrell, “Sun: A bayesian framework for saliency using natural statistics,” *J. Vision*, vol.8, no.7, pp.1–20, 2008.
- [23] X. Hou and L. Zhang, “Dynamic visual attention: Searching for coding length increments,” *Proc. Advances in Neural Information Processing Systems*, pp.681–688, 2008.
- [24] J. Harel, C. Koch, and P. Perona, “Graph-based visual saliency,” *Proc. Advances in Neural Information Processing Systems*, pp.681–688, 2007.

- [25] N. Bruce and J. Tsotsos, “Saliency, attention, and visual search: An information theoretic approach,” *J. Vision*, vol.9, no.3, pp.1–24, 2009.
- [26] X. Hou, J. Harel, and C. Koch, “Image signature: Highlighting sparse salient regions,” *PAMI, IEEE Trans.*, vol.34, no.1, pp.194–201, 2012.
- [27] M. Rubinstein, A. Shamir, and S. Avidan, “Multi-operator media retargeting,” *ACM Trans. Graph.*, vol.28, no.3, pp.23:1–23:11, July 2009.
- [28] 三柴 数, 池原雅章, “シームカービングを用いたリサイズ可能な画像の圧縮,” *信学技報*, vol.110, pp.43–48, 2010.
- [29] 三柴 数, 池原雅章, “ブロックベースシームカービングを用いた画像のリサイズ,” *信学誌*, vol.J93-A, pp.605–612, 2010.
- [30] D.D. Conger, M. Kumar, and H. Radha, “Multi-seam carving via seamlets,” *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol.7870, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Feb. 2011.
- [31] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, “A comparative study of image retargeting,” *ACM Trans. Graph.*, vol.29, no.6, pp.160:1–160:10, Dec. 2010.
- [32] P.S. Heckbert, “Fundamentals of texture mapping and image warping,” *Technical Report*, University of California at Berkeley, 1989.
- [33] L. Montejano, “Some results about minkowski addition and difference,” *Mathematika*, vol.43, pp.265–273, 1996.
- [34] 杉原厚吉, 今井敏行, 畑口剛之, “図形のミンコフスキー和の逆演算は何か,” *信学技報*, vol.96, pp.33–40, 1996.
- [35] 杉原厚吉, 今井敏行, 畑口剛之, “図形の可逆なミンコフスキー和の提案,” *信学誌*, vol.J80-D-2, pp.2663–2670, 1997.
- [36] J. Eckstein and D.P. Bertsekas, “On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Math. Program.*, vol.55, no.3, pp.293–318, 1992.

- [37] D. Gabay and B. Mercier, “A dual algorithm for the solution of non linear variational problems via finite element approximation,” *Computers & Mathematics with Applications*, vol.2, pp.17–40, 1975.
- [38] M.V. Afonso, J. Bioucas-Dias, and M.A.T. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems,” *Image Processing, IEEE Trans. on*, vol.20, no.3, pp.681–695, 2011.
- [39] M. Yamagishi, S. Ono, and I. Yamada, “Three variants of alternating direction method of multipliers without certain inner iterations and their application to image super-resolution via sparse representation,” *Technical Report, IEICE*, pp.49–54, 2012.
- [40] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer New York, 2011.

発表論文

本研究に含まれる発表論文

学会論文（査読あり）

- (第3章) 野中 敬介, 宮田 高道, 羽鳥 好律, “任意形ディスプレイのための Image Retargeting 手法の提案,” 映像情報メディア学会論文誌 12月号, pp. J524-J532, 2012年12月.
- (第4章) 野中 敬介, 宮田 高道, 羽鳥 好律, “ROI 移動の大域最適化による任意形イメージリターゲティングの提案,” 電子情報通信学会論文誌 D (条件付採録).

国際会議（査読あり）

- (第4章) Keisuke Nonaka, Takamichi Miyata, Yoshinori Hatori, “Generalized Image Retargeting via Convex Optimization ”, IEEE International Conference on Image Processing, pp. 1066–1070, Melbourne, Australia, Sep. 2013.
- (第4章) Keisuke Nonaka, Takamichi Miyata, Yoshinori Hatori, “Image Retargeting for Arbitrary-shaped Display via Convexification of ROIs Positioning Problem, ” International Workshop on Advanced Image Technology, SPS-10, p. 787, Nagoya, Japan, Jan. 2013.
- (第4章) Keisuke Nonaka, Takamichi Miyata, Yoshinori Hatori, “Generalized Image Retargeting via Convexification of ROIs Positioning Problem, ” Future Trend of Nanodevices and Photonics, IEEE EDS WIMNACT-37, P-94, Tokyo, Japan, Feb. 2013.

国内学会発表

- (第3章) 野中 敬介, 宮田 高道, 酒井 善則, “画像の非矩形変換による遮蔽回避の一検討,” 映像メディア処理シンポジウム, I2-04, pp. 29–30, 静岡県伊豆市, 2008年10月 (ベストポスター賞受賞) .
- (第3章) 野中 敬介, 宮田 高道, 酒井 善則, “非矩形視聴領域における画像の最適提示手法に関する基礎検討,” 映像メディア処理シンポジウム, I2-15, pp. 47–48, 静岡県伊豆市, 2009年10月.
- (第3章) 野中 敬介, 宮田 高道, 羽鳥 好律, “繰り返し性に着目した Image Retargeting のための重要領域抽出,” 映像情報メディア学会年次大会, No. 10, pp. 13–14, 成蹊大学, 2011年8月.
- (第4章) 野中 敬介, 宮田 高道, 羽鳥 好律, “再構成時のパッチの重みを忠実に利用したエネルギー汎関数の最小化による Image Retargeting 手法の基礎検討,” 映像メディア処理シンポジウム, I4-01, pp. 57–58, 静岡県熱海市, 2011年10月.
- (第4章) 野中 敬介, 宮田 高道, 羽鳥 好律, “Image Retargeting における重要領域移動問題の移動可能領域限定による凸化に関する検討,” 映像メディア処理シンポジウム, I1-03, pp. 11–12, 静岡県熱海市, 2012年10月.
- (第4章) 野中 敬介, 宮田 高道, 羽鳥 好律, “再構成時のパッチ重要性の合計値に着目した Image Retargeting 手法の一検討,” 第40回画像電子学会年次大会, P-4, 早稲田大学, 2012年6月 (研究奨励賞 受賞) .
- (第3章) 古賀 理, 野中 敬介, 吉野 知伸, 羽鳥 好律, “背景の連続性に着目した任意形領域に対する Image Retargeting 手法の高速化に関する検討,” 画像電子学会第264回研究会, pp. 104–111, 高知工科大学, 2013年3月.
- (第5章) 野中 敬介, 宮田 高道, 羽鳥 好律, “複数領域の効果的表示のためのマルチイメージリターゲットングに関する基礎検討,” 映像メディア処理シンポジウム, I3-02, 静岡県熱海市, 2013年11月 (ベストポスター賞受賞) .

付録I

I.1 定理4.4.3の証明

式 (4.5) における関数 ζ は下半連続で proper な凸関数であり, A_i , $W_{i,j}$, H_f , H_b および H_l は空でない閉凸集合である. そこで, 定理4.4.3を示すためには ζ の強圧性³を示せばよい (文献 [40] Section 11.4 参照). これを背理法を用いて証明する. すべての $m \in \mathbb{N}$ に対して点列 $(\mathbf{u}_m)_{m \in \mathbb{N}} \in \mathbb{R}^{2n}$ が存在し,

$$\|\mathbf{u}_m\|_2 \rightarrow +\infty, \quad (\text{I.1})$$

$$\lim_{\|\mathbf{u}_m\|_2 \rightarrow +\infty} \zeta(\mathbf{u}_m) \not\rightarrow +\infty. \quad (\text{I.2})$$

であると仮定する. そのとき, $\alpha \in \mathbb{R}_{++} := \{\xi \in \mathbb{R} \mid \xi > 0\}$, および以下の点列 $(\mathbf{u}_m^{\text{sub}})_{m \in \mathbb{N}} \in \mathbb{R}^{2n}$ が存在する.

$$\zeta(\mathbf{u}_m^{\text{sub}}) \leq \alpha \quad (\forall m \in \mathbb{N}). \quad (\text{I.3})$$

関数 ζ に含まれるすべての関数は, すべての $\mathbf{u} \in \mathbb{R}^{2n}$ に対して 0 以上を返すため, 不等式 (I.3) からすべての $m \in \mathbb{N}$ に対して

$$\iota_{H_l}(\mathbf{u}_m^{\text{sub}}) \leq \alpha \quad (\forall m \in \mathbb{N}), \quad (\text{I.4})$$

であるといえる. 指示関数の定義から

$$\mathbf{u}_m^{\text{sub}} \in H_l \quad (\forall m \in \mathbb{N}), \quad (\text{I.5})$$

である. H_l の有界性と (I.5) から, 明らかに,

$$\sqrt{2n} \leq \|\mathbf{u}_m^{\text{sub}}\|_2 \leq \sqrt{2n}w_T \quad (\forall m \in \mathbb{N}), \quad (\text{I.6})$$

³関数 $f: \mathbb{R}^N \rightarrow (-\infty + \infty]$ が $\lim_{\|\mathbf{x}\| \rightarrow +\infty} f(\mathbf{x}) = +\infty$ であるとする. このとき f は強圧的であるという.

または

$$\sqrt{2n} \leq \|\mathbf{u}_m^{sub}\|_2 \leq \sqrt{2nh_T} \ (\forall m \in \mathbb{N}), \quad (\text{I.7})$$

であることが保証される．これら式 (I.6) および式 (I.7) は (I.1) に反する．よって，関数 ζ は強圧的である．

I.2 定理 4.4.4 の証明

行列 U は単位行列であるため，式 (4.28) における行列 M は明らかに列フルランクである．そのため， $M^T M$ は正則である．式 (4.26) と (4.27) の対応関係により， $\text{dom}(\psi) = \mathbb{R}^{N_1}$ であることがわかり，これは行列 M の列フルランク性と同時に $0 \in \text{ri}(\text{dom}(\tilde{\psi}) - M\text{dom}(\psi))$ が成り立つことを暗に示している．したがって，問題 (4.5) は定理 4.4.2 の条件を満たしており， $\mathcal{L}_0(\mathbf{x}, \mathbf{y}, \mathbf{z})$ の鞍点が存在する．よって，定理 4.4.1 により Algorithm 4 の収束性が保証される．