焦点ボケ量推定を用いた単一画像からの 3次元空間における領域抽出方法

正員 杉浦 彰彦 (豊橋技術科学大学大学院)

学生員 鎌田 光宣 (豊橋技術科学大学大学院)

The Extracting Method from a Single Picture at 3-D Space by Using Defocusing Estimation

Akihiko Sugiura, Member, Mitsunobu Kamata, Student Member (Graduate School of Engineering,
Toyohashi University of Technology)

The diverse broadcast means that will be available in the future will cause an increased demand for programs. When the input of the posture of an agent is used to manipulate a virtual computer graphics actor, it is better if the system does not require a special studio and devices.

In the present paper, we propose a way to extract images from a single picture based on estimates of blooming. This is done using a partial auto-correlation analysis that carries out backward and forward predictions simultaneously. And, we divide targets into the stratified depth from a single image. An experiment was conducted using a picture taken with a digital camera, and satisfactory results were obtained.

キーワード：焦点ボケ, 単一画像, 領域抽出, 偏自己相関分析

1. ま え が き

今後、メディアの大部分がデジタルへと移行してゆくと、コンテンツの質の低下および枯渇が問題になることが予想される。例えば、TVがデジタル放送になり多チャンネル化が進むと、番組は個々の視聴者のニーズに応えるため、より一層の多様性が求められる。番組が増えると1つ1つの番組制作にかけられる予算が減り、コンテンツの質を高く保つだけの機材や人材の確保が困難となる。このような問題を解消するには、番組制作の簡素化が必要となってくる。

また、インターネット上で情報を発信する場合、現在ではサーバの処理能力や通信速度の面から文字と静止画が中心のWebページが主流であるが、今後、通信インフラが整備されることにより、音や映像をふんだんに使った情報を発信できるようになる。さらに、インターネット上では個人で情報発信を行う場合も多く、手軽にマルチメディアコンテンツを制作できる環境が求められる。

これらの要求に応えるため、1人でも簡単にコンテンツを制作できるパソコンソフト「ワンマンハリウッド」の開発を目指す。姿勢や動作を入力する方法として代表的なものに、接触デバイスを用いて磁気あるいは光学センサによりエージェントの動きを読み取るモーションキャプチャがある。また、接触デバイスを用いないものでは、ステレオカメラあるいは複数台のカメラから読み込んだデータから

3次元データを復元する方法が成果をあげている⁽¹⁾⁽²⁾。

しかしながら、手軽に扱えるようにするためには、特殊なスタジオセットや接触型のデバイスは使用しないことが望ましく、カメラも一般的なものを利用したい。また、ヒューマンインタフェースとして被写体の動きを解析する場合、おおまかな動きを抽出できれば十分である。そこで本プロジェクトでは人物の姿勢を1台のカメラから入力する方法を取る。

単眼視の画像から人物の姿勢データを抽出する際に、人体を少数のパーツに分割した人体スティックモデルを用いる手法がある⁽³⁾。これは、色情報やフレーム間差分で抽出したシルエット状の人物領域と、人体スティックモデルの姿勢をパターンマッチングすることで姿勢を推定する手法である。シルエットの抽出として抽出の手法としてよく利用されているものにクロマキー⁽⁴⁾やフレーム間差分を利用する方法などがある。しかしながら、これらの方法で得られた2次元の抽出領域から推定できる動きパターンは非常に限られてしまう。

本稿では人物領域の特定に、焦点ボケ量推定を用いた単一画像の距離測定方式⁽⁵⁾を応用して、単一画像から焦点の合っている部分を抽出する手法を提案する。本手法では単眼視の画像から人物の領域とともに奥行き方向のデータが得られるため、抽出できる姿勢の種類を増やすことができる。

図 1 焦点ボケ画像の生成機構
Fig. 1. The generation model of blooming picture.

2. 原 理

カメラで撮影した画像を見ると、焦点から外れた部分は高周波成分が低減された形で写っている。

図 1 に示すように、物体から出た光はレンズにより集光され、写像面に結像する。このとき、焦点が外れていると写像面の前後に結像するため、非収束光が重なり合った形で焦点ボケ画像となる⁽⁶⁾。焦点ボケ画像の生成を周波数次元で考えると、焦点ボケ信号は、元画像信号にレンズ系の点広がり関数を畳み込み積分した形で表現される。

点広がり関数のモデルを図 2 に示す。焦点部分では原点のみで振幅値が最大となり、焦点から遠くなるにつれて広がってゆく。そのため、入力された画像から焦点ボケ画像を生成するのと逆の手順を踏み、点広がり関数を求めることができれば焦点からの距離を求めることができる。ここで、本研究では点広がり関数の広がり具合を焦点ボケ量と呼ぶこととする。

本手法では、前向き予測と後ろ向き予測を同時に行う偏自己相関 (PARCOR: Partial auto-correlation) 分析法により、分析区間全体で共通に畳み込まれている成分を線形予測することで点広がり特性を推定する。画像データを小領域に区切り、小領域ごとに偏自己相関係数を求めて焦点ボケ量を推定する。

偏自己相関係数は次式で定義される⁽⁷⁾。

$$k_{n+1} = \frac{\overline{\epsilon_{f,t} \cdot \epsilon_{b,t-(n+1)}}}{\sqrt{\epsilon_{f,t}^2} \cdot \sqrt{\epsilon_{b,t-(n+1)}^2}} \dots\dots\dots (1)$$

図 2 点広がり関数のモデル
Fig. 2. The model of point spread function.

図 3 偏自己相関の定義
Fig. 3. Definition of partial autocorrelation.

式 (1) は図 3 に示すように、 x_t から $x_{t-(n+1)}$ の間の n 個のサンプルにおける前向き予測残差と後ろ向き予測残差間の相関であり、 x_t と n だけ離れたサンプル間の偏自己相関を表している。

ここで、 $\epsilon_{f,t}$ は実測値 x_t と予測値 x'_t との誤差であり、 x_t の前向きに予測できる成分を除いた前向き予測残差を示す。また、 $\epsilon_{b,t-(n+1)}$ は後ろ向きに予測できる成分を除いた後ろ向き予測残差を示す。 x_t の前向き予測係数を $a_i (i = 1, \dots, p)$ 、後ろ向き予測係数を $b_i (i = 1, \dots, p)$ とすれば、

$$\epsilon_{f,t} = x_t - x'_t = x_t + \sum_{i=1}^n a_i x_{t-i} = \sum_{i=0}^n a_i x_{t-i}$$

$$a_0 \equiv 1 \dots\dots\dots (2)$$

$$\epsilon_{b,t-(n+1)} = x_{t-(n+1)} - x'_{t-(n+1)} = x_{t-(n+1)} + \sum_{j=1}^n b_j x_{t-j} = \sum_{j=1}^{n+1} b_j x_{t-j}$$

$$b_{n+1} \equiv 1 \dots\dots\dots (3)$$

となる。 x_t が一定なら、 $b_j = a_{n+1-j} (j = 1, \dots, n+1)$ である。

このように定義された偏自己相関係数を、Durbin-Levinson-Itakura 法を用いて求める。この手法は線形予測分析に広く使われているもので、 $n(n+1)$ 回の積和と n 回の除算で n 次までの線形予測係数と偏自己相関係数が得られる。

偏自己相関係数は焦点付近では低次数にパワーが集中し、焦点から遠ざかるにつれて広がってゆく。そこで、 i 次の偏自己相関係数を $f(x_i)$ としたとき、焦点ボケ量 B は式 (4) に示すように N 次までの偏自己相関係数の 2 乗和とする。求めた焦点ボケ量から、画像の中で焦点ボケ量の小さな部分、すなわち焦点の合っている領域を求めて抽出する。

$$B = \sum_{i=1}^N f(x_i)^2 \dots\dots\dots (4)$$

なお、カメラや被写体が動くことによっても入力画像にボケが生じるが、動きボケ自体はシャッタースピードを上げることで抑えることができる。また、信号を1次元で考えると、動きボケが後方に広がるのに対して焦点ボケは前後方向に広がることになり、それぞれ独立した関数で表現される。そのため、本手法で求める焦点ボケ量の相対的な大きさには影響しない。

3. 抽出処理

入力画像から、焦点距離にある領域を抽出する処理の手順を以下に述べる。

まず、フルカラーで入力された画像データを256段階の輝度値に変換する。

次にノイズの除去を行う。偏自己相関分析を行うにあたり、入力データはフィルタを通さずに扱うほうが望ましいが、そのままではノイズのピクセルを焦点距離にあるものと判断してしまうため、ホワイトノイズを取り除くフィルタが必要となる。

ノイズ除去フィルタとして、平均化法、k最近傍法、メディアン法、エッジ保存平滑化法を試した⁽⁸⁾⁽⁹⁾。この中で、平均化法、k最近傍法のぼかしフィルタでは、焦点距離にあるエッジ部分にボケが生じてしまい、焦点ボケ量の推定が不可能になってしまう。実験の結果、本手法においては、解像度を低下させずにノイズを除去するメディアンフィルタが最適であった。

そして次に、ノイズを除去した画像データを使って焦点ボケ量を求める。偏自己相関分析の1区間を256個とし、画像を16×16ピクセルのブロックに分割して焦点ボケ量を推定する。ブロックのデータを縦方向、横方向でそれぞれ1次元配列に格納し、その際に生じる16ピクセルごと

の境界をなくすため、奇数列では左から右へ、偶数列では右から左へと読み取る。

1次元配列に格納されたデータはオフセットカットフィルタを通して偏自己相関分析を行い、偏自己相関係数を得る。なお、焦点ボケ量推定の精度を上げるため、ブロックを数ピクセルずつスライドさせて計算を行い、各ピクセルの焦点ボケ量は平均値をとって求める。

焦点ボケ量が求まったら、画面内で最小の焦点ボケ量を0、最大の焦点ボケ量を255として正規化し、しきい値以上のピクセルを抜き出す。このとき、画面上にはしきい値以上の領域としきい値以下の領域が複数存在している状態である。そこで、しきい値以上の領域のうちサイズが最大のものを抽出対象とみなし、その領域の内側に含まれるしきい値以下の領域を抽出対象に含め、逆に外側にあるしきい値以上の領域は抽出対象から除外する処理を行うことにより、1つの連続した領域が出力される。

画像により抽出に最適なしきい値は異なるが、上述の処理では焦点ボケ量のしきい値をあらかじめ手動で設定しておく必要がある。そこで、しきい値を自動で変更する方法を以下に述べる。しきい値はあらかじめ低く設定しておく。

手順1. 焦点ボケ量がしきい値以上のピクセルを抽出領域として抜き出す。

手順2. 抽出対象となる領域は画像内に点在している状態である。抽出領域全体と最大の領域とのサイズを比較し、サイズ差が大きければしきい値を増やして手順1から繰り返す。(図4)

手順3. 最大領域の内側と外側に分け、それぞれの領域を1つにまとめる。

手順4. 手順2で求めた領域のサイズと手順3で求めた領域との排他的論理和をとり、その結果がある値以上であればしきい値を増やして手順1から繰り返す。(図4)

画像中の抽出対象の領域に対して抽出領域ははじめ小さく、ほぼ一致するまで領域を広げていくことで、自動的に最適なしきい値を求める。連続した枚数の抽出を行う際は、直前に得たしきい値を用いることで収束までの回数を減らすことができる。

4. 合焦点にある対象の抽出

デジタルカメラから入力した単画像より、合焦点(ピントの合っている部分)に位置する対象を抽出する実験を行う。抽出したい物体に焦点を合わせると、背景の部分には焦点ボケが生じる。そこで、画像中で焦点ボケの少ない領域を抜き出すことで対象を抽出する。本手法では焦点ボケの推定に偏自己相関分析を用いるが、この方法で対象の領域が抽出できることを確かめる。

実験には人物写真、風景写真、大小2種類の白黒のチェッカー模様を用意した。このうち小さい方のチェッカー模様は人物の形に切り抜き、大きい方のチェッカー模様は背景として使う。また、クロマキーとの比較を行うため、背景に青いカーテンを用意した。(図5)

図4 しきい値自動設定の概要

Fig. 4. Summary of automatic threshold setting.

図 5 用意した抽出対象と背景
Fig. 5. Prepared targets and backgrounds.

本手法では背景と物体を分離しやすくするために、被写界深度を浅くする必要がある。しかしながら、焦点距離を長くすると視野角が狭くなってしまい、また広い視野角を保ったまま焦点距離を長くするにはレンズの口径を大きくする必要がある。この問題を解決するための方法としてスクリーン投影形撮像系⁽¹⁰⁾が考案されているが、今回はそのような特殊な装置は使用せず、狭い視野角に収まる被写体を用いることとした。また、被写体が奥行き方向に広がりを持っていると、被写体の領域内でも焦点ボケの大きさが変化してしまう。今回の実験では、抽出処理をより容易に行うため、奥行き方向に広がりを持たない被写体を用いることとした。

まず、抽出対象として人物写真あるいは切り抜いたチェッカー模様と焦点を合わせる。今回用意した2つのチェッカー模様、すなわち抽出対象となる切抜きの模様と背景の模様のサイズ差は2.8倍であった。そこで、背景を抽出対象の2.8倍の距離に設置した。こうすることで、カメラには同じ大きさで焦点ボケ量の異なる模様が入力されることになり、背景と同じテクスチャでも問題なく領域を抽出できることが確認できる。今回は抽出対象をカメラから60cmの距離に設置し、背景はそこから約2.8倍にあたる170cmの距離に設置した。(表1)(図6)

表 1 実験諸元

Table 1. Condition of experiment system.

入力デバイス	総画素数: 約 150 万
デジタルスチルカメラ	圧縮形式: 非圧縮
処理画素	640 × 480 画素中 画面中央部分 約 350 × 400 画素
絞り値	ワイド端 2~11 テレ端 2.4~11
レンズ	5.2mm~26mm (35mm フィルムカメラ換算 28mm~140mm)
撮影距離	抽出対象 60cm 背景 170cm

これらをデジタルカメラから入力し、抽出対象の写っている画面中央部分を切り抜いて抽出処理を行う。なお、入力される画像に JPEG などの圧縮がかけられていると、高周波成分のカットやマクロブロックが生じるなどして焦点ボケ量の推定に支障を来すため、非圧縮のデータを用いた。また、ボケ量推定時に求める偏自己相関係数の次数を決める必要があるが、今回の実験画像においては次数を変えても抽出結果にほとんど影響がなく、5 次まで求めることとした。

推定された焦点ボケ量と抽出結果の画像を図7に示す。焦点ボケ量を表す画像では、各ピクセルの焦点ボケ量を 256 段階の輝度で表しており、輝度が高いほど焦点ボケ量が大きくなっている。また、抽出結果には2種類あり、ひとつはしきい値を手動で設定したものであり、もうひとつは前述の方法によりしきい値を自動で設定したものである。この場合のしきい値とは、求めた焦点ボケ量より、どの値で被写体と背景とを分けるかを定める値である。また、手動で設定したときの結果は、抽出結果を見ながらしきい値を

図 6 配置図
Fig. 6. Figure of arrangement.

図 7 実験結果
Fig. 7. Extraction results.

表 2 実験結果
Table 2. Extraction results.

しきい値を手動で設定						
人物	背景	しきい値	実サイズ [Pixels]	抽出サイズ [Pixels]	ミスヒット数 [Pixels]	ミスヒット率
グレー	グレー	37.1%	43117	50699	10654	24.7%
グレー	チェッカー	39.1%	42808	43519	5295	12.4%
チェッカー	グレー	32.4%	43479	52789	9766	22.5%
チェッカー	チェッカー	26.2%	42905	44348	4833	11.3%
しきい値を自動で設定						
人物	背景	しきい値	実サイズ [Pixels]	抽出サイズ [Pixels]	ミスヒット数 [Pixels]	ミスヒット率
グレー	グレー	34.0%	43117	41534	14415	33.4%
グレー	チェッカー	39.1%	42808	42336	5578	13.0%
チェッカー	グレー	32.4%	43479	56648	13429	30.1%
チェッカー	チェッカー	26.2%	42905	43723	4838	11.3%
クロマキー						
人物	背景	しきい値	実サイズ [Pixels]	抽出サイズ [Pixels]	ミスヒット数 [Pixels]	ミスヒット率
グレー	青	—	42046	39810	2280	5.4%

変えてゆき、見た目に最適であると判断したときのしきい値を用いている。抽出結果の画像を見ると、実際の対象領域よりも大きい領域が抽出される傾向にあることが分かる。

そして、抽出時の焦点ボケ量のしきい値とミスヒット率を表 2 に示す。表中にある実サイズとは、画面を見ながら手動で入力画像のボケ領域を塗り潰し、残った部分すなわち対象の写っている領域のピクセル数である。ミスヒット数とは、手動で入力画像のボケ領域を塗り潰した画像と抽出結果とを比較して、異なるピクセル数をカウントしたものである。次式に示すように、一方では背景領域と判定されているのにもう片方では抽出領域となっている場合、またはその逆の場合にカウントされる。また、ミスヒット率とは、ミスヒット数を元サイズとの比率で表したものである。

表は左から人物と背景のタイプ、抽出時の焦点ボケ量のしきい値、比較元の人物サイズ、抽出サイズ、ミスヒット数、ミスヒット率となっている。クロマキーのミスヒット率が 5.4% と最も低く、背景がチェッカーの場合は 11% ～ 13%、背景が写真の場合は 20% ～ 30% のミスヒット率となった。

5. 離散コサイン変換を用いた手法との比較

偏自己相関分析を用いる本手法と、従来手法の 1 つである離散コサイン変換を用いた抽出との比較を行う。離散コサイン変換により画像の持つ周波数成分を調べてボケ量を計算し、ボケ量の少ない領域を合焦点付近にあるものとして抽出する。2 つの手法を比較するため、含まれる周波数成分の分布が異なる画像を用意してテクスチャへの依存の程度を調べる。

ここではボケ量を式 (5) として定義する。図 8 に示すように、 m, n は周波数領域全体のサイズを表し、 k, l は低周波領域のサイズを表している。これにより、低周波域にエネルギーが集中することでボケ量が大きくなる。今回用いた入力画像では低周波域の範囲を変更しても抽出結果への影響がほとんどなかったため、 k, l をそれぞれ m, n の 1/2

図 8 DCT 適用後の低周波域
Fig. 8. Low frequency area after DCT.

図 9 用意したチェッカー模様の切抜き
Fig. 9. Prepared cutout doll.

表 3 模様のサイズ		
Table 3. Pattern Size.		
	背景の模様 [mm]	抽出対象の模様 [mm]
A	15	7
B	22	11
C	33	16

とした。なお、本手法と比較するため、入力画像を 16 × 16 ピクセルのブロックに区切って離散コサイン変換を行った。

$$B = \frac{\sum_{i=0}^k \sum_{j=0}^l (a_{i,j})^2}{\sum_{g=0}^m \sum_{h=0}^n (a_{g,h})^2} \dots\dots\dots (5)$$

$$m = n = 16, k = \frac{m}{2}, l = \frac{n}{2}$$

被写体として、図 9 に示す 3 種類のチェッカー模様の切抜きを用意した。切抜きの形状とサイズは同一であり、模様のみ大きさが異なる。また、この切抜きに対応するチェッカー模様の背景を 3 種類用意し、サイズ順に A、B、C の組に分けた。表 3 に示す数字はチェッカー模様の白い四角の一辺の長さであり、どの組も切抜きと背景の模様のサイズは約 2 倍になっている。入力画像上で背景と切抜きの模様が同じ大きさになるようにするため、カメラから背景までの距離を切抜きまでの距離の 2 倍になるよう設置した。

求めたボケ量を用いて被写体の抽出処理を行うわけだが、2 つの手法を比較するために、抽出結果のミスヒット率を被写体の内側と外側に分けて求め、抽出対象である内側のミスヒット率が 2% 以下になるまでしきい値を変化させて抽出した。なお、しきい値の初期値は最大ボケ量の 50% である 128 とした。

このときのしきい値とミスヒット率の関係を図 10 に示す。グラフの中で、中抜きの記号は被写体領域内側のミスヒット率を表し、塗り潰されている記号は外側のミスヒット率を表している。また、本手法である偏自己相関 (PARCOR) のグラフではしきい値の範囲が 125 から 155 まで表示してあるのに対し、離散コサイン変換 (DCT) のグラフでは 125 から 245 までの範囲を表示している。

図 10 模様の大さによるミスヒット率の変化
Fig. 10. Relation of pattern size and miss-hit rate.

どちらの手法とも、内側と外側のミスヒット率にはトレードオフの関係があるが、初期値として用いたしきい値 (128) の付近では本手法のミスヒット率が低い。また、離散コサイン変換を用いた手法では内側のミスヒット率を減らすためにしきい値を大きく変える必要があり、その規模は画像の持つ周波数成分の特性により異なる。自動で処理を行う場合、画像の特徴を調べてしきい値を変更することは困難である。このことから、偏自己相関分析を用いる本手法の方がテクスチャへの依存が小さいと言える。

6. 対象の奥行き方向への振り分け

求めた焦点ボケ量を用いて、画像から物体の奥行き方向の情報を得る。カメラから背景に向かって空間を複数のレイヤに区切り、対象とする物体が奥行き方向のどの位置にあるかを推定してレイヤに振り分ける実験を行う。これにはカメラから対象までの距離と焦点ボケ量の関係が明らかになっている必要があり、はじめに距離と焦点ボケ量の関係を求める実験を行う。また、画像から対象を抽出するには対象が背景からある程度離れている必要があるが、その距離も実験により求める。

まず、カメラから被写体までの距離と焦点ボケ量の関係を明らかにする。焦点ボケ量に大きく影響する被写界深度は、被写体までの距離の他、カメラの各種パラメータによって大きく変わり、焦点ボケ量と距離をそのまま結びつけることはできない。被写界深度は計算により求めることがで

表 4 後方被写界深度

Table 4. Depth of backward field[mm].

Distance to a target [mm]	f = 15	f = 20	f = 26
400	12.4	7.1	4.2
800	48.1	27.8	16.7

きるため、この値をレイヤ分割の基準に利用することを考える。そこで、被写体の距離とカメラの焦点距離を変えて撮影した画像から焦点ボケ量を求め、被写界深度との関係を調べる。

実験に用いたデジタルカメラは許容錯乱円 $\delta = 7.5 \times 10^{-3}$ [mm] であり、 F_{NO} の値を 2.4 として撮影した。被写体を $L = 400$ [mm] と $L = 800$ [mm] の位置に配置し、焦点距離は $f = 15$ [mm], $f = 20$ [mm], $f = 26$ [mm] の 3 通りに設定して撮影した。これらを下式に当てはめ、求めた後方被写界深度の値を表 4 に示す

後方被写界深度

$$d_b = \frac{\delta F_{NO} L^2}{f^2 + \delta \cdot F_{NO} L} \dots\dots\dots (6)$$

実験には大小 2 種類のチェッカー模様を用意した。被写体として長方形に切った小さなチェッカー模様を用い、背景として大きいチェッカー模様を用いた。

カメラから 400mm または 800mm の距離にある被写体に焦点を合わせ、背景となるチェッカー模様を 10mm ずつ遠ざけて撮影する。この入力画像から焦点ボケ量を計算し、背景部分の焦点ボケ量と被写体部分の焦点ボケ量の比を求めた。被写体から背景までの距離と焦点ボケ量の比の関係を図 11 に示す。横軸には抽出対象と背景との距離、縦軸は焦点ボケ量の比を示している。グラフを見ると、後方被写界深度と焦点ボケ量の比の間に相関があることが分かる。

続いて、この結果に基づき奥行きを複数のレイヤに分ける実験を行う。実験には 3 種類のチェッカー模様を用意する。ひとつは、背景として後方被写界深度の 10 倍の距離

図 11 距離と焦点ボケ量の関係
Fig. 11. Relation of distance and Blooming Volume.

図 12 距離と各レイヤ含有率との関係

Fig. 12. Relation of distance and component rate of each layer.

に設置する。2つ目は合焦点に設置するもので、カメラから 400mm または 800mm の距離に設置し、カメラの焦点をそこに合わせる。そしてもうひとつは、抽出対象として合焦点と背景の間を 10mm 刻みで動かす

これらを先ほどの実験と同じように、被写体距離（合焦点）を $L = 400[\text{mm}]$, $L = 800[\text{mm}]$, 焦点距離を $f = 15[\text{mm}]$, $f = 20[\text{mm}]$, $f = 26[\text{mm}]$ として撮影した。そして、合焦部分の焦点ボケ量を 0%, 背景部分の焦点ボケ量を 100% として、抽出対象を焦点ボケ量に応じて 3 つのレイヤに振り分ける。焦点ボケ量が 25% 以下の場合にレイヤ 1 とし、これは前述の実験より後方被写界深度の約 2.5 倍となっている。同じく 25% から 75% まではレイヤ 2 とし、75%（後方被写界深度の約 5 倍）以上をレイヤ 3 とした。

図 12 に、距離とレイヤ含有率の関係を示す レイヤ含有率とは、画像中で抽出対象の写っている領域の焦点ボケ量を求め、各ピクセルがどのレイヤに含まれるかを調べてその比率を求めたものである。背景が後方被写界深度の 10 倍の距離に設置されているため、グラフの一番右が後方被写界深度の 10 倍の距離になっている。グラフを見ると、後方被写界深度の約 5 倍で被写体と背景をほぼ完全に分離できることが分かる。また、被写体を 3 つの奥行き方向のレイヤに分けることも可能であることが分かる。

7. む す び

合焦点に位置する物体を抽出する実験では、単画像から

偏自己相関分析を用いて焦点ボケ量の推定が可能であることを示し、また最適なしきい値を自動で設定する方法を述べた。被写体抽出の実験では、背景がチェッカー模様ではほぼ期待通りの結果が得られ、そして風景の写真でも良好な結果が得られた。よって、偏自己相関分析を用いて焦点ボケ量を推定する本手法が、単一画像における被写体の抽出に有効なことが確認できた。

続いて、偏自己相関を用いる本手法と離散コサイン変換を用いた手法とを比較するため、それぞれの手法で画像のボケ量を推定し、しきい値を変えながら被写体を抽出する実験を行った。実験の結果、離散コサイン変換では画像の持つ周波数成分の特性によりしきい値を大きく変える必要があることが分かった。一般に、自動で処理を行う場合、画像の特徴を調べてしきい値を変更することは困難であるため、偏自己相関分析を用いる本手法の方がテクスチャへの依存が小さいと言える。

最後にレイヤ分割の実験を行った。はじめに距離と焦点ボケの関係性を求め、その結果を基に、被写体が奥行き方向のどこに位置するのかを推定して各レイヤに振り分けた。実験の結果、後方被写界深度と焦点ボケ量の比の間に相関があることが分かった。被写界深度は計算により求めることができるため、この値をレイヤ分割の際の基準とすることができる。今回の実験では、カメラから後方被写界深度の 10 倍の距離までの間を 3 つのレイヤに区切ることができた。

これを人物の姿勢抽出に応用することを考える。例えば人物の右腕の部分カメラ側に突き出すと、右腕の領域の焦点ボケ量が変化するため、その領域がカメラ側に突き出されたと判定することができる。このとき、撮影系の被写界深度の理論値と認識可能な層の数、あるいは抽出精度との関係を求める。これにより得られた情報から人物の姿勢を 3 次元の奥行き情報を持つデータとして抽出し、他の形式に変換して保存したり、別のキャラクターに変換して表示することが可能となる。

今後は得られた人物領域をもとに、単一画像から人物の姿勢データを抽出することを試みる。

（平成 12 年 7 月 26 日受付，同 13 年 4 月 2 日再受付）

文 献

- (1) 岩沢昭一郎・海老原一之・竹松克浩・坂口竜乙・大谷淳: "Shall We Dance?" の構築. 信学技報 PRMU98-114, pp.15-22 (1998)
- (2) 岡松利美・大橋健・乃間司・松尾英明・江島俊郎: 奥行き情報を利用した人物追跡システム ステレオ Ptracker. 信学技報 PRMU99-210, pp.45-52 (2000)
- (3) 木本伊彦・梶谷昭彦・安田康彦: スティックモデルに基づく単眼視動画像からの人体歩行運動の解析の一手法. 信学論 D-II, Vol. J74-D-II, No.3, pp.376-387 (1991)
- (4) 日本放送協会編: テレビジョン番組制作技術. 日本放送出版協会 (1983)
- (5) Akihiko Sugiura, Kiyoharu Aizawa and Hiroshi Harashima: Ranging a Blooming Picture by Using Partial Autocorrelation, Electronics and Communications in Japan, Part 2, Vol.82, No.1, pp.23-33 (1999)

(6) 藤波重次：写真技術 改訂版，共立出版 (1963)
(7) 中田和男・南敏：音声・画像工学，昭見堂 (1987)
(8) M. Nagao and T. Matsuyama：Edge Preserving Smoothing, Computer Graphics and Image Processing 9, pp.394-407 (1979)
(9) 森俊二・坂倉梅子：特徴抽出，エッジ検出，テクスチャー解析，(画像認識の基礎:2) オーム社 (1990)
(10) 大原知之・三輪道雄・石井雅博・小池康晴・佐藤誠：合焦法距離計測のためのスクリーン投影型画像観測系の提案，信学論 Vol.J82-D-2, No.10, pp.1750-1757 (1999)

杉 浦 彰 彦 (正員) 1965 年 10 月 16 日生まれ，平成 2 年東京農工大学大学院修士課程了，同年豊田工業大学制御情報工学科 助手，現在 豊橋技術科学大学大学院工学研究科 助教授，マルチメディア情報通信，高能率符号化，画像応用，スペクトル拡散，CDMA，Bluetooth，携帯電話方式，デジタル放送方式などの研究に従事，工博，電子情報通信学会，IEEE 会員。

鎌 田 光 宣 (学生員) 1976 年 9 月 17 日生まれ，平成 13 年豊橋技術科学大学大学院工学研究科修士課程修了，現在，同大学大学院工学研究科博士後期課程在学中，映像メディア情報の研究に従事，電子情報通信学会，映像情報メディア学会会員