

結像位置を変化させて得た多重画像からの パンフォーカス画像の自動合成法

Pan-Focused Image Synthesis Using Multiple
Images with Different Focal Conditions

塩 昭夫

Akio SHIO

NTTヒューマンインターフェース研究所

NTT Human Interface Laboratories

あらまし： 本報告では、結像位置を変えて得た複数の画像から画像の全ての部分に焦点の合った画像（パンフォーカス画像と呼ぶ）を合成する方法を提案する。厚みを持った物体を画像化すると通常はその一部にのみピントの合った画像になる。そこで、本方法では、まず（1）結像位置を少しづつ変えながら複数の画像を入力し、（2）各画像の部分領域毎にフォーカス度（画像の濃度の正規化標準偏差を使用）を求める、（3）フォーカス度が最大となる結像位置の部分領域画像を選択し、（4）これらをつなぎ合わせてパンフォーカス画像を合成する。本報告では、幾つかの実験結果からアルゴリズムの有効性を示す。

Abstract: This paper proposes a pan-focused image (all image components are in perfect focus) synthesis algorithm to combine multiple-images with different optical distances. An object with depth is usually transformed into an image some parts of which are defocused. In this algorithm, (1)multiple images are acquired with different optical distances, (2)a focus measure is evaluated for small regions of the multiple images with different focal conditions, where the focus measure is defined by the normalized standard deviation of the gray-level values in each region, (3)from the multiple images, regions in perfect focus are selected based on this measure, and (4)the pan-focused image is synthesized from the selected regions. The algorithm was confirmed by experiments using several samples.

Keywords: Image Processing, Image Synthesis, Focus, Pan-Focused Image, Auto-Focus, Depth from Focus, Multiple Images.

[1] まえがき

一般に、奥行きのある物体を画像化すると、物体の一部にのみピントのあった不鮮明なものになる。例えば、物体に近接した位置での画像入力や顕微鏡画像の入力などの場合、被写体焦点深度が浅いため、厚みのある物体では焦点ボケが起こる。

そこで筆者は、結像位置を変化させて撮影した複数の画像（以下では単に原画像と呼ぶ）からボケのない鮮明な画像（Pan-Focused Image:以下では P F 画像と

呼ぶ）を合成する方法を検討した。本技術は、物体に近接した位置で画像を入力する種々の対象に適用可能であり、画像処理の要である入力画像の品質向上に寄与するものである。

この研究に関連するものとしては、近年注目されている画像のボケを利用した奥行き検出法 (Depth from Focus)^{[1]-[6]} や、家庭用を含めた各種カメラに搭載されている焦点の自動調節機能^{[7]-[9]}などがあるが、P F 画像の合成に関する研究は見られない。本報告で

は、PF画像を合成するアルゴリズムを提案すると共に、幾つかの実験結果からその有効性を示す。

[2] アルゴリズム

本方法は、焦点深度を大きく越えるような奥行きのある物体を対象として、全ての部分に焦点の合った鮮明な画像を合成することを目的としている。まず、結像位置nを変えながらN枚の原画像 I_n ($n = 1, N$)を撮影する。次に、各原画像を格子状に分割して $p \times q$ 画素の部分領域rを考え、N枚の部分画像 i_{rn} ($n = 1, N$)のうち最もピントの合っている部分画像 i_{rn} を以下で述べる尺度(以下ではフォーカス度と呼ぶ)を基準に選択する。この様にして得られた部分画像をつなぎあわせ、1枚のPF画像を合成する。

この方法は、フォーカス度を利用する点で焦点ぼけを利用した奥行き検出法(Depth from Focus)^{[1]-[6]}およびオートフォーカス技術^{[7]-[9]}に類似している。しかし、前者は距離情報を求める目的としていること、後者に関しては、一点または平均的な合焦点位置を求める目的とする点で本手法のアプローチと異なっている。

2-1 原画像入力系の構成

図1は、原画像入力系の構成を示している。物体を横切る平面 d_n とセンサがレンズを介して結像関係にあるとすれば、レンズから平面 d_n およびセンサまでの距離をそれぞれ a_n 、 b_n 、レンズの焦点距離を f として、

$$1/a_n + 1/b_n = 1/f \quad (1)$$

が成り立つ。結像位置を変化させる方法としては、

[方法1] 距離 a_n 、 b_n を固定し、カメラ(センサ+レンズ)と物体の相対距離を変える、

[方法2] レンズで距離調節を行う(距離 b_n を調節して結像位置を変える)、

の2つが考えられる。

a. 方法1

方法1は、距離 a_n 以下の奥行きしかカバー出来ない反面、距離 a_n 、 b_n が変わらないため得られる原画像の倍率補正が不要であるという長所がある。焦点ぼけが問題になるのは、通常近距離の場合に限られていることから、方法1は、倍率補正の必要な方法2に比べ簡単さの点で有利である。

b. 方法2

方法2は、近距離から無限遠までの距離をカバーできる反面、レンズ位置の移動により画角が変化するため得られる画像の倍率が一定にならない問題がある。

任意の結像位置nに対するセンサ上の像の大きさを l_n 、大きさの基準とする結像位置n'に対する像の大きさを $l_{n'}$ とするとき、 l_n と $l_{n'}$ の間には

$$\begin{aligned} l_n &= R \cdot l_{n'} \\ &= \frac{a_n \cdot (a_{n'} - f)}{a_{n'} \cdot (a_n - f)} \cdot l_{n'}, \end{aligned} \quad (2)$$

の関係がある([付録]を参照)。距離 a_n 、 $a_{n'}$ が共に焦点距離fに比べて充分大きい時には、式(2)のRの値がほぼ1になるため原画像の大きさを補正する必要はないが、そうでない時には補正($1/R$ 倍する)が必要である。

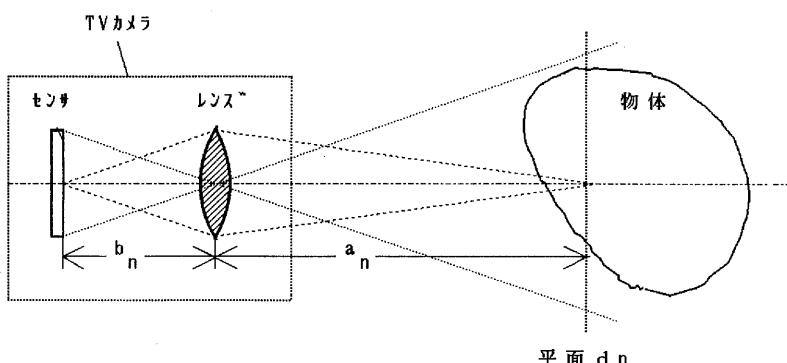


図1. 画像入力系の構成

2-2 フォーカス度

フォーカス度の決定は、任意の部分画像に関して焦点の合っている結像位置を選択するためのパラメータを決める問題であり、Depth from Focusやオートフォーカス技術と共に問題と見ることができる。

Depth from Focus^{[1][4][5]}に於いては、光学系の点像分布関数をガウス関数で近似し、エッジ部分のボケの空間分散から奥行きを算出するなどの方法を探っている。

一方、オートフォーカスの為の尺度に関しては、主に濃度値に関するパラメータから合焦点位置を検出している。Lighthart^[7]らの比較研究では、11種のパラメータ（濃度勾配の絶対値、ラプラシアン、標準偏差など）のうち濃度勾配の自乗および正規化標準偏差が良好な結果を得ている。

ボケの問題を正確に議論する上では、前者のように画像の空間的変化を調べるのが自然であるが、処理量から見ると、後者の濃度変化を用いる方法が有利である。何れの方法も、ボケによって濃度コントラストが減少すること利用している点で同じである。

物体から反射されてセンサ上に到達する光強度は、結像位置の変化に伴って変化するから、原画像の濃淡値も変化する。フォーカス度は、この様な強度変化の影響を受けないことが必要である。

カメラ系の特性を考慮した正規化標準偏差は、照明条件が変わっても近似的に物体本来の反射率の変化の度合を表す^[10]ことから、ここではフォーカス度として濃度の正規化標準偏差（Lighthart^[7]らの用いたものと類似しているが、カメラ特性を考慮している点で異なる）を用いる。

$$[定義] \quad C_{rn} = \sigma_{rn} / (\mu_{rn} - \beta) \quad (3)$$

ただし、 σ_{rn} と μ_{rn} は、それぞれ $p \times q$ 画素の部分画像 i_{rn} の濃度の標準偏差および平均、 β はカメラ系で決定されるパラメータである。

2-3 PF 画像の合成

上記の画像入力系により得られた原画像から、フォーカス度に基づいてPF画像を合成する方法について述べる。原画像（方法2では大きさを補正したもの）は、ピントの合った部分に関しては、オクルージョンがなければ物体上の同一部分を投影する。しかし、焦点のボケた部分では、画像の倍率が変わるために位置ずれを生じる。すなわち、任意の部分領域 r に対する N

個のフォーカス度 C_{rn} ($n = 1, \dots, N$) は、必ずしも物体上の同一部分に対応しない問題がある。また、濃度の一様な領域では、フォーカス度が不安定になる。

そこで、物体の奥行き変化が滑らかであると仮定し、得られたフォーカス度を平滑化した後に、最大フォーカス度をもつ結像位置を部分領域ごとに決定する方法をとる。

平滑化処理では、濃度の一様な領域のフォーカス度が不安定であることから、フォーカス度 C_{rn} をその隣接部分領域内の最大フォーカス度 C'_{rn}

$$C'_{rn} = \max_{j \in k} (C_{jn}) \quad (4)$$

(k は、部分領域 r の8近傍内の部分領域)
で置き換える処理：

$$C'_{rn} \rightarrow C_{jn} \quad (5)$$

を繰り返し行う。

〔3〕実験

3-1 データ

実験には、方法1および方法2による2種のデータを使用した。

a. データ1

対象物体（67°傾けて置いた鉛筆）上の結像位置を奥行き方向に3mm間隔で変えながら17枚の原画像を撮影した（方法1による）。その一例を図2(a)(b)(c)に示す。パラメータは以下の通り。

$$\begin{aligned} a_n &= 260.3\text{mm}, & b_n &= 69.7\text{mm}, \\ D(\text{開口径}) &= 9.8\text{mm}, & f &= 55\text{mm}, \\ F(\text{絞り}) &= 5.6, & p = q &= 8. \end{aligned}$$

b. データ2

対象物体（約45°傾けて置いた英和辞典）の各部分に関し、焦点の合った原画像が得られる様に、レンズのピントを適宜に調節しながら4枚の原画像を撮影した（方法2による）。その例を図3(a)(b)(c)に示す。パラメータは以下の通り。

$$\begin{aligned} a_n &= 292.2 \sim 323.7\text{mm} \\ b_n &= 67.8 \sim 66.3\text{mm} \\ (D, f, F, p, q) & \text{は、データ1の場合と同じ} \end{aligned}$$

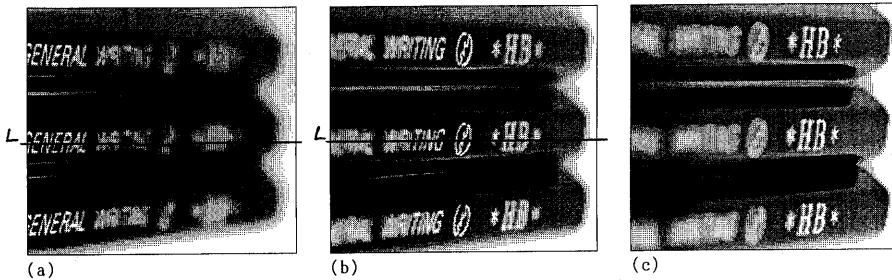


図 2. 原画像(1)の例

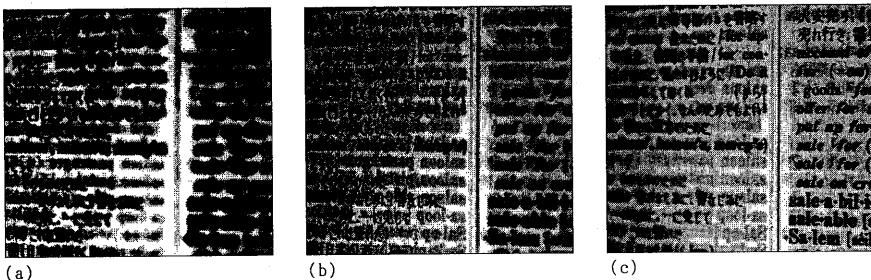


図 3. 原画像(2)の例

3-2 結果

a. データ1に対する結果

図4は、図2(a)(b)の直線Lにそった位置(横軸)に対するフォーカス度の値(縦軸)を示し、「■」「○」は平滑化する前、「+」「×」はそれらの平滑化後の結果である。平滑化する前のフォーカス度はかなり不安定であるが、平滑化によって安定化しているのが見られる。

次に、この平滑化されたフォーカス度から、各部分領域に対して正しい結像位置の原画像が選択されることを示す。図5は、直線L上にある部分領域に対するフォーカス度から得られた距離(奥行き)の対応付け結果であり、横軸は直線L上の位置座標、縦軸は相対距離(奥行き)を示す。これから、各部分領域とも良好な結像位置が選択されていることが判る。

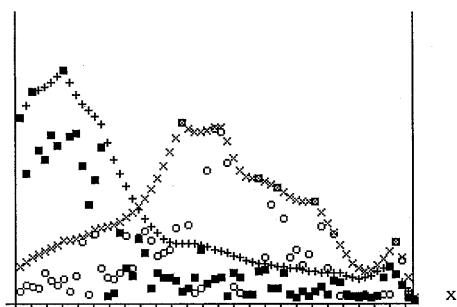
図6に、データ1から合成されたPF画像を示す。文字列の近傍に若干の雑音が見られるが、ほぼ満足で

きる結果である。また、方法1で原画像を入力するとカメラと物体上の結像位置までの距離が一定(倍率も一定)になるため、合成されたPF画像(図6)は、実際にはかなり斜め(67°)の方向から見ているにも関わらず、正面から見ているように感じられる(遠近感がない)。

b. データ2に対する結果

図7は、データ2から合成されたPF画像である。2-1b節で述べた様に、方法2で原画像を入力する場合には一般には倍率補正が必要になる。しかし、この例では、奥行きの変化が連続的であること、および最も離れた部分領域間でも倍率差が10%程度であり、隣接部分領域間では0.2%程度になることから倍率補正是行っていない。得られたPF画像(図7)は、自然であり、各部分領域間の継目はほとんど判らない。

フォーカス度



■:図2(a)に対するフォーカス度, +:■の平滑化結果
○:図2(b)に対するフォーカス度, ×:○の平滑化結果

図4. フォーカス度の観測値

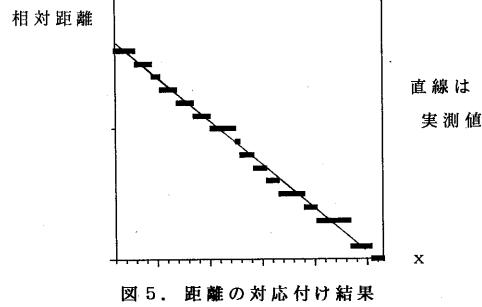


図5. 距離の対応付け結果

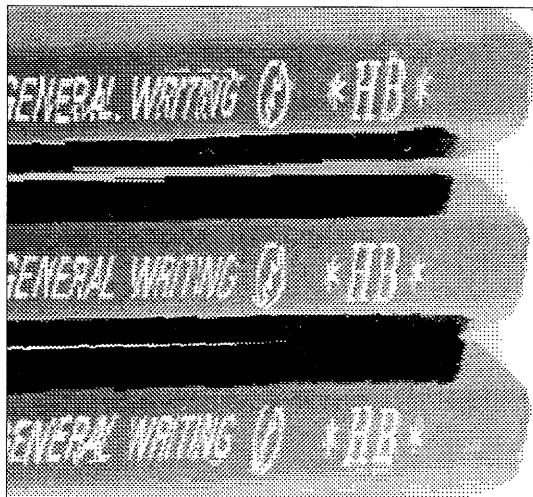


図6. データ1から合成されたPF画像

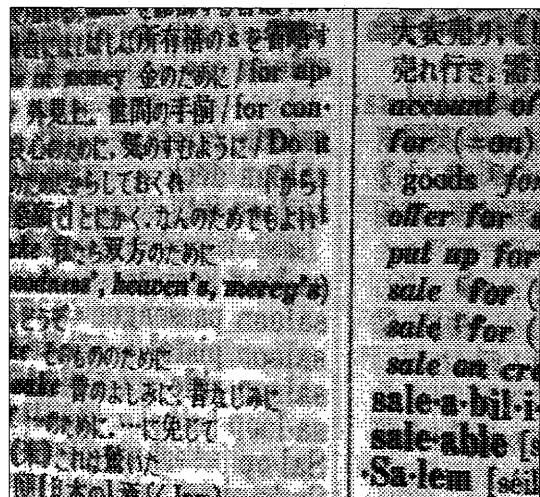


図7. データ2から合成されたPF画像

[4] むすび

結像位置を変えながら取り込んだ複数の原画像をもとに、全ての部分に焦点の合った画像（PF画像）を自動的に合成する方法について述べた。

この方法は、原画像の濃度情報を利用しており、対象物体に奥行きなどの情報は必要としていない。もし、各原画像に対する結像位置（奥行き）が既知の場合には、安野らの方法^[6]と同様に奥行きの計測にも利用可能である。

ここでは、フォーカス度として、カメラ系を考慮した正規化標準偏差を用いることによって、自然照明下の対象にも適用可能とした。本報告では、2種の撮像光学系について検討し、その両者に対して良好なPF

画像合成結果を得た。

今後の課題としては、物体の奥行き変化量と原画像の入力間隔の関係を明らかにすると共に、種々の物体を対象に本手法を評価することなどである。

[謝辞] 本研究の機会を与えて戴いた小森和昭視覚情報研究部長、同部酒井高志グループリーダ、石井健一郎主幹研究員、日頃ディスカッション戴く奥平雅士主幹研究員はじめグループ各位に感謝致します。

[参考文献]

- [1] A.P. Pentland, "A new sense for depth of field," IEEE Int. J. Conf. Artificial Intelligence, (IJCAI'85) pp. 988-994, 1985.

- [2] E. Krotkov and J. Martin, "Range from focus," IEEE Int. Conf. Robotics and Automation (ICRA'86), vol. 2, pp. 1093-1098, 1986.
- [3] P. Grossmann, "Depth from focus," Pat. Recog. Letters 5, pp. 63-69, 1987.
- [4] T. Darrell, K. Wohn, "Pyramid based depth from focus," IEEE Comput. Soc. Conf. Comput. Vision and Pat. Recog. (CVPR'88), pp. 504-509, 1988.
- [5] M. Subbarao and N. Gurumoorthy, "Depth recovery from blurred edge," IEEE Comput. Soc. Conf. Comput. Vision and Pat. Recog. (CVPR'88), pp. 498-503, 1988.
- [6] 安野, 岡田, 横山, 北川, "連続的に焦点距離が変化する多重画像を用いた立体計測," 情処第36全大 5W-8, pp. 1875-1876, 1988.
- [7] G. Ligthart and C. A. Groen, "A comparison of different autofocus algorithms," IEEE Int. Conf. Pat. Recog. (ICPR'82), pp. 597-600, 1982.
- [8] 山本, 田宗, 田村, "距離画像の入力と処理," 信学技報 PRU86-129, pp. 49-59, 1966.
- [9] 坂野, "オートフォーカス技術," 光学, vol. 12, no. 5, pp. 351-358, 1983.
- [10] 塩, "情景中文字の検出のための動的二値化処理法" 信学論 D, vol. J71-D, no. 5, pp. 863-873, 1988.

【付録】：補正倍率の算出

図 A は、原画像の倍率補正に関する説明図であり、理解を容易にするためレンズ中心の位置を固定した状態で、レンズ中心から等角度にある大きさ L_n , $L_{n'}$ なる 2 つの物体の像（大きさ l_n , $l_{n'}$ ）がレンズにより形成されているものとする。

まず、2 つの像の大きさは、

$$l_n = L_n / (b_n / a_n) \quad (A 1)$$

$$l_{n'} = L_{n'} / (b_{n'} / a_{n'}) \quad (A 2)$$

であり、2 つの物体が等角度の位置にあることから、

$$L_n / a_n = L_{n'} / a_{n'} \quad (A 3)$$

が成り立つ。

a_n と b_n , $a_{n'}$ と $b_{n'}$ は、それぞれ独立でないから、(1) 式を用いて (A 1), (A 2) 式の b_n , $b_{n'}$ を消去すると、

$$l_n = L_n f / (a_n - f) \quad (A 4)$$

$$l_{n'} = L_{n'} f / (a_{n'} - f) \quad (A 5)$$

となる。更に、式 (A 3) により L_n を消去すると、式 (A 5) は、

$$l_{n'} = a_n L_n f / a_{n'} (a_{n'} - f) \quad (A 6)$$

となる。

従って、(A 4) 式と (A 5) 式から、2 つの像の大きさの比 R :

$$\begin{aligned} R &= l_n / l_{n'} \\ &= a_n (a_{n'} - f) / a_{n'} (a_n - f) \end{aligned} \quad (A 4)$$

が得られる。すなわち、像の大きさ l_n は、 $l_{n'}$ の R 倍になっている。

従って、結像位置 n' に対する像の大きさ $l_{n'}$ を基準にして原画像の倍率を一定にする為には、各原画像の大きさ l_n を $1 / R$ 倍すればよいことになる。

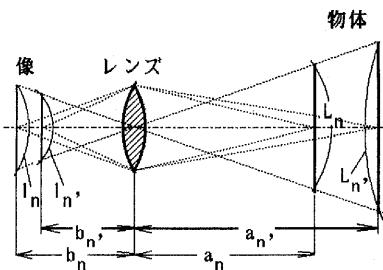


図 A. 原画像の倍率