

高品質な映像生成に適した焦点ぼけ画像群の構造化に関する検討

児玉 和也[†] 久保田 彰^{††}

[†] 情報・システム研究機構 国立情報学研究所

〒101-8430 東京都千代田区一ツ橋 2-1-2

^{††} 東京工業大学大学院 総合理工学研究科

〒226-8502 横浜市緑区長津田町 4259-G2-31

E-mail: [†]kazuya@nii.ac.jp, ^{††}kubota@ip.titech.ac.jp

あらまし 筆者らは、焦点合わせの異なる多数枚の撮像画像群を3次元情報とみなし、これから様々な映像を生成する手法を検討している。本稿ではまず、撮像画像間の撮像範囲の相違や焦点合わせ間隔の補正等により、焦点ぼけ画像群を適切に構造化する手法を検討する。あわせて、構造化後の3次元焦点ぼけを表すパラメータと生成画像の品質との関係を雑音への耐性も含めて検証し、望ましい撮像条件を明らかにする。また、画像生成時に仮定したぼけの程度や対象物体の存在範囲などからの誤差の影響による品質劣化についても検証を行い、本手法が有効に機能する範囲を示すとともに、窓関数を用いた劣化抑制等の必要性の有無など一般の劣化画像の復元処理との相違についても論ずる。キーワード 焦点ぼけ、3次元、全焦点画像、映像生成、周波数領域

A Study on Structure of Multi-Focus Imaging Sequences for High Quality Image Generation

Kazuya KODAMA[†] and Akira KUBOTA^{††}

[†] National Institute of Informatics, Research Organization of Information and Systems

Hitotsubashi 2-1-2, Chiyoda-ku, Tokyo, 101-8430 Japan

^{††} Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

Nagatsuta 4259-G2-31, Midori-ku, Yokohama, 226-8502 Japan

E-mail: [†]kazuya@nii.ac.jp, ^{††}kubota@ip.titech.ac.jp

Abstract This paper deals with image generation based on transformation integrating multiple differently focused images into structured 3-D information. First, we describes a method of image size correction and focus interpolation for the integration. In addition, relation between a certain parameter of the structured 3-D information and quality of generated images is discussed taking noise into consideration. Furthermore, we examine how errors of assumptions about blurs and scenes degrade generated images. Difference between our proposed method and conventional image restoration such as necessity of using window functions is also mentioned.

Key words blur, 3-D, all-in-focus image, image generation, frequency domain

1. はじめに

筆者らは、焦点合わせの異なる多数枚の撮像画像群 [1]-[13] を3次元情報とみなし、これから様々な映像を生成する手法を検討している [12], [13]。幾何光学的なぼけモデルから導き出される3次元フィルタを用いれば、対象シーンと撮像画像群を空間位置不変の畳み込み演算により結びつけることができる。これに基づき例えば全焦点画像を、シーン推定を介さず直接に撮像画像群から生成するフィルタが設計可能である。

本稿ではまず、撮像画像間の撮像範囲の相違や焦点合わせ間隔の補正等により、焦点ぼけ画像群を適切に構造化する手法を検討する。あわせて、構造化後の3次元焦点ぼけを表すパラメータと生成画像の品質との関係を雑音への耐性も含めて検証し、望ましい撮像条件を明らかにする。また、画像生成時に仮定したぼけの程度や対象物体の存在範囲などからの誤差の影響による品質劣化についても検証を行い、本手法が有効に機能する範囲を示すとともに、窓関数を用いた劣化抑制等の必要性の有無など一般の劣化画像の復元処理との相違についても論ずる。

2. 多数枚のぼけ画像群からの映像生成

2.1 幾何光学モデルに基づくぼけ画像群の構造化

本手法では、図1のようにまず理想的な幾何光学モデルによる撮像を仮定する[12]。撮像面 P_i で合焦していた対象が撮像面 P_j において生じるぼけ半径 r_{ij} を、各撮像面のレンズ面からの距離 v_i, v_j およびレンズ半径 L を用いて表せば、

$$r_{ij} = \frac{|v_i - v_j|}{v_i} L \quad (1)$$

となる。ここで、ある基準撮像面 P_b に合わせ面 P_j での取得画像のサイズ補正を行なった場合、補正後のぼけ半径 $\bar{r}_{ij}$ は、基準撮像面のレンズ面からの距離を v_b とすれば、

$$\bar{r}_{ij} = \frac{v_b}{v_j} r_{ij} = \frac{v_b}{v_j} \frac{|v_i - v_j|}{v_i} L \quad (2)$$

と書ける。逆に面 P_j で合焦していた対象が面 P_i において生じるぼけ半径は、

$$\bar{r}_{ji} = \frac{v_b}{v_i} r_{ji} = \frac{v_b}{v_i} \frac{|v_i - v_j|}{v_j} L \quad (3)$$

と表せるので $\bar{r}_{ji} = \bar{r}_{ij}$ の関係が導ける。また、同様にして3つの撮像面 P_i, P_j, P_k (レンズ面からの距離は $v_i \geq v_j \geq v_k$ とする) の間で互いに生じるぼけ半径は、

$$\bar{r}_{ik} = \frac{v_b}{v_k} \frac{v_i - v_k}{v_i} L \quad (4)$$

$$\bar{r}_{ij} = \frac{v_b}{v_j} \frac{v_i - v_j}{v_i} L \quad (5)$$

$$\bar{r}_{jk} = \frac{v_b}{v_k} \frac{v_j - v_k}{v_j} L \quad (6)$$

であることから、 $\bar{r}_{ik} = \bar{r}_{ij} + \bar{r}_{jk}$ の関係が成立する。

以上から、 N 個の撮像面 $P_0 \sim P_{N-1}$ を適切に調整すれば、適当な $r(\geq 0)$ を用いて、

$$\bar{r}_{ij} = |j - i| r \quad (7)$$

なる条件を満たす、構造化されたぼけ画像群の撮像が可能である。

2.2 3次元フィルタによる空間周波数の構造化分析

前節で示した条件に基づいて構造化された撮像画像群にサイズ補正を施したものを、画像平面 (x, y) に垂直な方向 (z 軸とする) に、撮像面 $P_0 \sim P_{N-1}$ (レンズ面からの距離は $v_0 > v_1 > \dots > v_{N-1}$ と設定する) に対応する順に $z = 0, 1, \dots, N-1$ と並べた3次元情報を $g(x, y, z)$ とおく。

また、各撮像画像について、対応する撮像面で合焦する対象奥行き平面にのみ注目し、その奥行き平面から得られる輝度値成分だけを取り出した画像群を新たに仮定する。これらに $g(x, y, z)$ と同様のサイズ補正を施したものを同一順に並べ $f(x, y, z)$ と定義すれば、対象シーンの3次元情報を表すことができる。すなわち $f(x, y, z)$ は、対象シーンからオクルージョン領域を除き、さらに適当な幾何変換を行ったものとなっている。

以上で定めた $f(x, y, z)$ と $g(x, y, z)$ は、前節で示した撮像条件から、3次元ぼけフィルタ $h(x, y, z)$ を用いた畳み込み演算により図2のように結びつけることができる。ただし、いま分

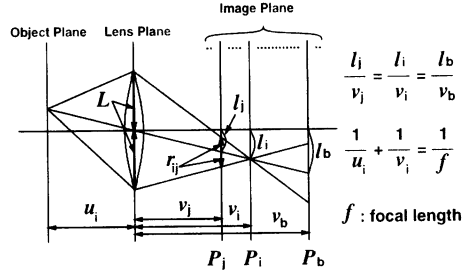


図1 幾何光学的なぼけモデル

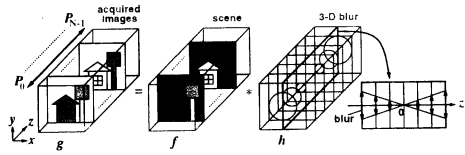


図2 3次元ぼけフィルタによる畳み込みの概念図

散 σ^2 で表されるガウスぼけを $b(x, y, \sigma)$ と表し、半径 R の幾何光学ぼけを $\sigma = R/\sqrt{2}$ により置き換えれば、

$$h(x, y, z; r) = b(x, y, r/\sqrt{2}) \quad (8)$$

である (ただし、 $\lim_{\sigma \rightarrow 0} b = \delta$) [12]。図2の畳み込み演算は周波数領域上で、

$$G(u, v, w) = H(u, v, w; r) F(u, v, w) \quad (9)$$

と変換されることから、 $H(u, v, w; r)$ の特性を分析すれば、撮像画像群 $g(x, y, z)$ が対象シーン $f(x, y, z)$ から、どのような空間周波数成分をどれくらい取り出すことになるのを知ることができる。

2.3 3次元フィルタを用いた映像生成

前節の3次元ぼけフィルタ $h(x, y, z)$ は、 $b(x, \sigma)$ を分散 σ^2 の1次元ガウス関数とすれば周波数領域上で、

$$H(u, v, w; r) = Nb(w; r(K_x^2 u^2 + K_y^2 v^2)^{1/2} / \sqrt{2}) \quad (10)$$

と変換される。ただし、補正後の各画像サイズを $N_x \times N_y$ としたとき、 $K_x = N/N_x, K_y = N/N_y$ である。ここで改めて $H(r) = H(u, v, w; r)$ と書くものとして、画像サイズが 128×128 で $N = 64$ とした場合の $H(1.0), H(0.0)$ の特性を図3に示す。どちらも (u, v) に関して楕円対称であり、 w 軸方向に、分散のみ異なるガウス関数かその分散0の極限としてのデルタ関数を描いていることがみてとれる。

$H(1.0)$ では $(u, v) = (0, 0)$ すなわち画像平面方向の直流成分について、 w 軸方向に $H(u, v, w) = 0 (w \neq 0)$ とデルタ関数を描く。これは対象シーン中の一様なテクスチャ領域については奥行き情報が取り出せないことを意味する。

また、 $H(0.0)$ は理想的なピンホールカメラで取得した撮像画像群に対応する3次元フィルタを表す。この場合、各撮像画像は全て同一の全焦点画像となるため、 $H(u, v, w)$ はデルタ関

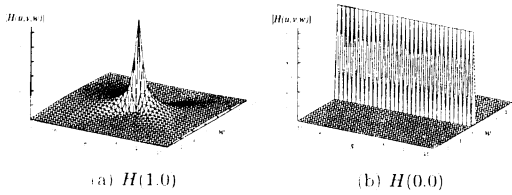


図3 3次元ぼけフィルタ $h(x, y, z)$ の周波数特性

数のみで構成される。これは対象シーンのテクスチャ情報は完全に取得可能だが、奥行き方向に関する一切の情報が失われることを意味する。

対象シーンの空間周波数情報をできる限り取得するという観点からみると、 $|H(r)|$ が一様にある程度大きいほど望ましい。 $H(1.0)$ を詳細に検討すると、 (u, v) が原点から離れるほど、すなわち画像平面方向の高周波成分ほど $|H(1.0)|$ は w 軸方向に一様に1に近付くことがわかる。これはシーン中のテクスチャの鮮明な領域については、奥行き方向の情報が詳細に取得できることを表す。

さて、 $w = 0$ すなわち (u, v) 平面の周波数成分に注目すると、 $|H(0.0)|$ に比較し $|H(1.0)|$ は小さいが、十分に1以上の利得が得られている。したがって一般に、次式により撮像画像群から、対象シーンの推定を介さず、直接に所望の全焦点画像 $a(x, y, z)$ へ変換することが考えられる。

$$\begin{aligned} A(u, v, w) &= H(0.0)F(u, v, w) \\ &= H(0.0)H^{-1}(r)G(u, v, w) \end{aligned} \quad (11)$$

上記の $H(0.0)H^{-1}(r)$ を新たなフィルタとみて特性を考察すると、 $H(r)$ の零点は $H(0.0)$ により打ち消されるので、全焦点画像の生成に必要な成分を全て撮像画像群から取り出せると期待できる。

3. 全焦点画像生成の実験

3.1 シミュレーション実験

シミュレーションとして、図4に示す様々な奥行きを持つシーンを仮定し図5のような64枚のぼけ画像群を合成した。図5(a)と図5(b)はそれぞれ $r = 1.0$ と $r = 0.5$ により構造化されている。これらに対し、 $H(0.0)H^{-1}(r)$ のフィルタを適用することで図6のような全焦点画像を生成した。所望の鮮明な画像が得られており、 $H(r)$ に含まれる零点近辺の影響を受けていないことが確認できる。

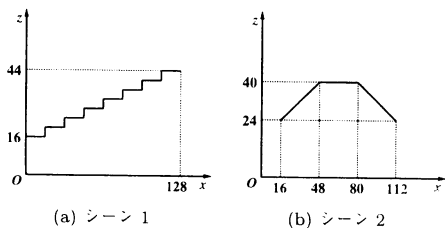


図4 シーンに仮定された奥行き (y は任意)

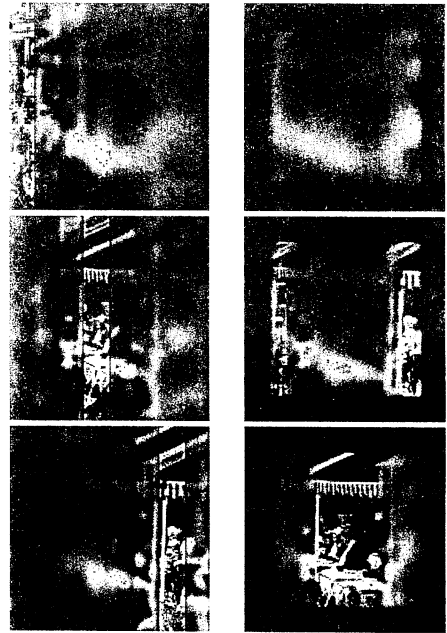


図5 合成した撮像画像 (128×128 画素) 群 $g(x, y, z)$
(上から $z = 16, 28, 40$ のみ示した)



図6 生成された全焦点画像 $a(x, y, z)$ (z は任意)

3.2 実画像を用いた実験

図7(a)-(c)に示すような連続的に焦点合わせを変化させながら撮影した64枚の実画像を用いて同様の実験を行う。各画像は 512×480 画素で構成されている。

図7(a)-(c)は、前章で述べた条件に焦点ぼけ画像群を構造化するため、既提案手法[8]に基づき、実際に取得された画像に対しサイズ補正と焦点合わせ間隔の補間処理を適用した後の画像である。構造化のパラメータは $r = 0.41$ と推定された。

カメラ特性が既知であり、カメラパラメータが精度良く調整できるならばこのような前処理は極めて容易に実現できる。この場合、焦点合わせ間隔の補間処理は必要なく r もカメラパラメータから決定できる。サイズ補正も推定処理を介することなく同様に決定できる。

図7(d)-(f)に従来の選択的統合法[7]と提案手法の画像生成結果を示す。図7(d)は従来手法の鮮鋭判定結果である。白い領域が遠景、黒い領域が近景として判定されたことを表す。この結果に基づき、鮮鋭領域を合成して全焦点画像を再構成した

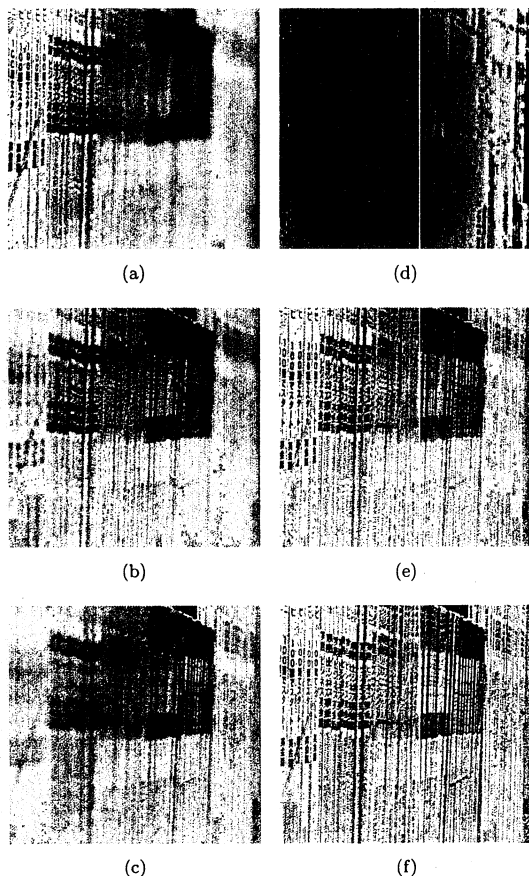
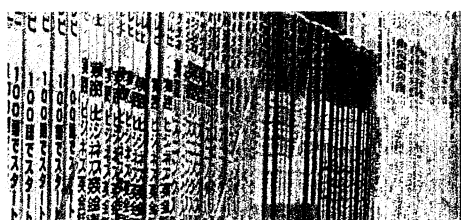
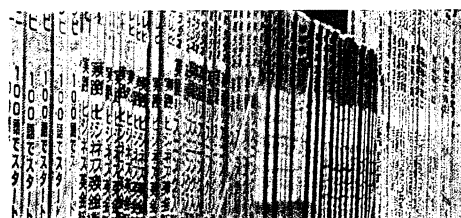


図7 実画像からの全焦点画像の生成: (a) $g(x, y, 16)$, (b) $g(x, y, 28)$, (c) $g(x, y, 40)$, (d) 鮮鋭判定結果, (e) 従来の選択的統合手法, (f) 提案手法



(a) 従来の選択的統合手法



(b) 3次元フィルタリングに基づく提案手法

図8 生成画像の上側半分の領域の拡大図

結果が図7(e)である。図7(f)が提案手法による全焦点画像の生成結果である。

図8に、各手法による生成画像の上半分の領域を拡大して示す。従来手法については、図7(d)における不安定な判定結果に起因する画質劣化が確認できる。これに対し提案手法では、画面全体にわたって鮮鋭な画像が比較的安定に生成されている。

4. 焦点ぼけ画像群からの生成画像の品質

以下では、焦点ぼけ画像群からの生成画像の品質へ影響しうる様々な要因を検証する。

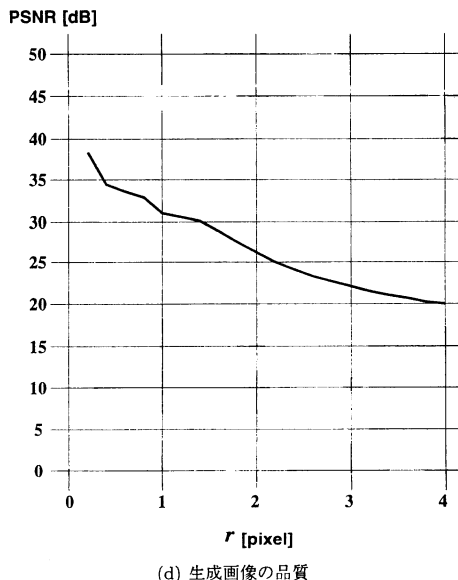
4.1 映像生成の際の仮定からの誤差の影響

焦点ぼけ画像群からの映像生成においては、対象シーン中の全ての領域がいずれかの撮像画像で合焦していることが望ましい。しかるに実際には焦点合わせ間隔が離散的であるため、もっとも合焦している撮像画像においても、焦点合わせがわずかにずれておりやや焦点ぼけしてしまう領域が存在しうる。

そこで、 $z = 31.5$ 上に存在する対象シーンを $z = 0, \dots, 63$ に焦点を合わせて撮像したと仮定し、 r を様々に変化させながら焦点ぼけ画像群を合成した。この画像群から生成した全焦点画像およびその品質を図9に示す。 r が大きくなるほど、もっとも焦点の合った画像でもある程度ぼけた画像となってしまうため生成画像の品質が劣化していくのがみてとれる。視覚的には $r \leq 1.0$ 程度が撮像条件として望ましいことが確認された。

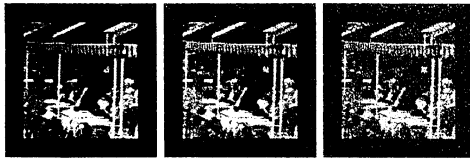


(a) $r = 0.0$ (b) $r = 2.0$ (c) $r = 4.0$

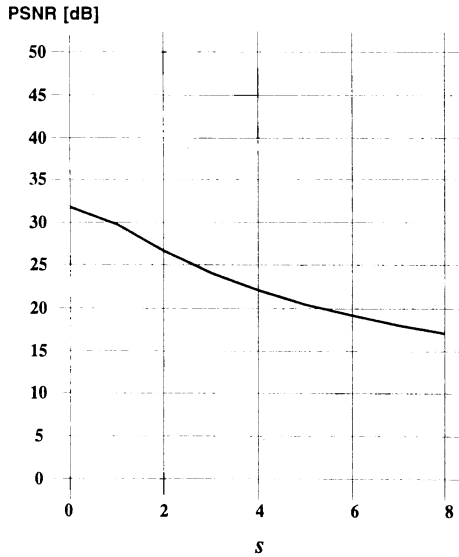


(d) 生成画像の品質

図9 対象シーンの連続的な奥行き分布に起因する品質への影響



(a) $s = 0.0$ (b) $s = 4.0$ (c) $s = 8.0$



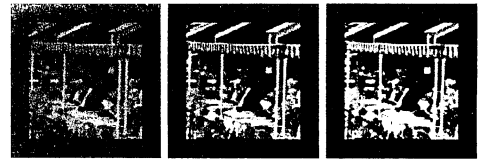
(d) 生成画像の品質

図 10 撮像画像群中の雑音に起因する品質への影響

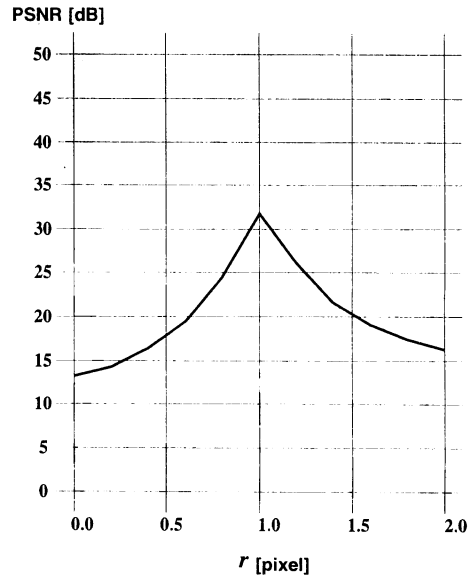
次に、 $r = 1.0$ として撮像画像群における雑音の影響を検証する。一般にはぼけによる劣化画像からの復元処理においては、雑音の生成画像への影響は極めて大きい。そこで、分散 s^2 (画像の輝度値は 0 から 255 とする) の白色雑音を加えた焦点ぼけ画像群から生成した全焦点画像およびその品質を図 10 に示す。

生成画像からみてとれるように、雑音の影響はあるがその領域は画素単位で限定的であり、一般のぼけ画像からの復元処理のような広範囲のアートファクトは生じていない。本提案手法の 3 次元フィルタの安定性によるものと考えられる。品質も雑音が強くなるほど悪くなるが、その変化は比較的滑らかであり急激な減少はみられない。

ところで、焦点ぼけ画像群が適切に構造化されておらず、例えば映像生成の際に仮定するパラメータ r の 3 次元焦点ぼけから何らかの誤差を含んでいた場合には、生成画像の品質が劣化するものと考えられる。具体的には、 r の大きさに関する誤差や、3 次元焦点ぼけそのものの歪みも考えられる。ただし、画像サイズの変化は画像サイズそのものに比して極めて小さいので、 r に関する誤差と比較すれば 3 次元焦点ぼけの歪みは無視できる。そこで、 $r = 1.0$ で構造化された焦点ぼけ画像群から、様々な r を仮定して生成した全焦点画像およびその品質を図 11 に示す。 $r < 1.0$ の場合にはコントラスト不足の画像が生成され、 $r > 1.0$ の場合にはコントラストが過剰に強調される。視覚的には、仮定した r に 2 割程度の誤差があっても比較的良好



(a) $r = 0.4$ (b) $r = 1.0$ (c) $r = 1.6$



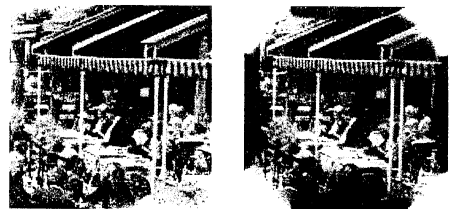
(d) 生成画像の品質

図 11 仮定したぼけパラメータ r との誤差の品質への影響

な全焦点画像が得られた。以上により、映像生成に際し焦点ぼけ画像群は仮定通りの 3 次元焦点ぼけで構造化されているほど良いが、ある程度の誤差を含んでいても視覚的には有意な生成画像が得られると考えられる。

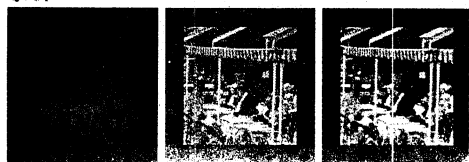
4.2 対象物体の存在範囲の品質への影響

一般のぼけ画像からの復元問題の場合、画像の端の領域で劣化モデルからのずれが生じそのままでは復元画像の品質が低下してしまうため、窓関数を適用してから復元を行う場合が多い。これに対し、本提案手法では対象シーンは焦点を合わせた範囲のはば中央に存在することを仮定しており、また 3 次元情報そのものの復元を行うのではないため、窓関数の適用が必ずしも有効とは考えられない。



(a) 窓関数の適用なし (b) 窓関数の適用あり

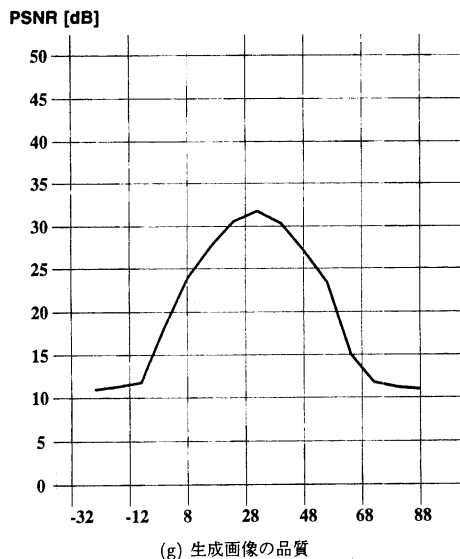
図 12 窓関数の適用の有無による生成画像の品質の変化



(a) $z = -16$ (b) $z = 0$ (c) $z = 16$



(d) $z = 48$ (e) $z = 64$ (f) $z = 80$



(g) 生成画像の品質

図 13 対象シーンの奥行きによる品質への影響

そこでまず、対象シーンの奥行きが画像の端まで一様として焦点ぼけ画像群を合成し、これから窓関数の適用の有無で条件をかえて全焦点画像の生成を行った。生成画像の比較を図 12 に示す。窓関数を適用しなくても、画像端周辺にわずかな影響はあるが一般の復元問題と異なり比較的画面上全体に鮮鋭な画像が得られている。窓関数を適用することによる品質の劣化を考慮すれば、本提案手法においてはとくに窓関数を利用することが有効とは考えられない。

次に、対象シーンの奥行きが仮定された焦点を合わせた範囲のはば中央に存在しない場合の影響を検証する。なお、奥行き方向に関しては遮蔽により一意に対象シーンが定まるので、窓関数はあきらかに不要である。ここでは、対象シーンの奥行き z を一定とし、これを様々に変更して焦点ぼけ画像群を合成し、これから全焦点画像の生成を行った。その生成結果と品質を図 13 に示す。仮定通り焦点を合わせた範囲のはば中央 ($z = 32$) に対象シーンが存在する場合、もっとも品質が高くなるが、 $z = 0$ から $z = 63$ の範囲に焦点を合わせたとすれば、例えば $z = 16$

から $z = 47$ に存在する対象シーンについては良好な品質の全焦点画像が得られていることがみてとれる。このことから本提案手法は、対象シーンが多様な奥行きを持つ場合にも、その奥行きの推定を介することなく直接に全焦点画像を生成できることがわかる。

5. おわりに

本稿では、焦点合わせの異なる多数枚の撮像画像群を 3 次元情報とみなしこれから様々な映像を生成する手法について、まず、撮像画像間の撮像範囲の相違や焦点合わせ間隔の補正等により、焦点ぼけ画像群を適切に構造化する手法を検討した。あわせて、構造化後の 3 次元焦点ぼけを表すパラメータと生成画像の品質との関係を雑音への耐性も含めて検証し、望ましい撮像条件を明らかにした。また、画像生成時に仮定したぼけの程度や対象物体の存在範囲などからの誤差の影響による品質劣化についても検証を行い、本手法が有効に機能する範囲を示した。

文 献

- [1] M.Subbarao, et al.: "Focused Image Recovery from Two Defocused Images Recorded with Different Camera Settings," IEEE Transactions on Image Processing, Vol.4, No.12, pp.1613-1627, 1995.
- [2] M.I.Sezan, et al.: "Survey of Recent Developments in Digital Image Restoration," Optical Engineering, Vol.29, No.4, pp.393-404, 1990.
- [3] P.J.Burt, et al.: "Enhanced Image Capture Through Fusion," Proc. 4th ICCV, pp.173-182, 1993.
- [4] P.J.Burt: "A Gradient Pyramid Basis for Pattern-Selective Image Fusion," Proc. SID 1992, pp.467-470, 1992.
- [5] M.Subbarao, et al.: "Application of Spatial-Domain Convolution/Deconvolution Transform for Determining Distance from Image Defocus," Computer Vision Laboratory, Stony Brook, Tech.Report 92.01.18, 1992.
- [6] 児玉他: "反復法に基づく複数画像からの任意焦点画像の生成", 映像学誌, Vol.51, No.12, pp.2072-2081, 1997.
- [7] K.Kodama, K.Aizawa, and M.Hatori: "Generation of arbitrarily focused images by using multiple differently focused images," SPIE Journal of Electronic Imaging, Vol.7, No.1, pp.138-144, 1998.
- [8] A.Kubota, K.Kodama, and K.Aizawa: "Registration and blur estimation methods for multiple differently focused images," 1999 IEEE International Conference on Image Processing, Vol.II, pp.447-451, 1999.
- [9] A.Kubota, and K.Aizawa: "Inverse filters for reconstruction of arbitrarily focused images from two differently focused images," 2000 IEEE International Conference on Image Processing, Vol.I, pp.101-104, 2000.
- [10] K.Aizawa, K.Kodama, and A.Kubota: "Producing Object-Based Special Effects by Fusing Multiple Differently Focused Images," IEEE Transactions on Circuits and Systems for Video Technology, Vol.10, No.2, pp.323-330, 2000.
- [11] 椿他: "多数の昆虫顕微鏡画像からの全焦点画像と 3 次元画像の生成", 映像学誌, Vol.56, No.2, pp.263-270, 2002.
- [12] 児玉: "多数の焦点合わせ画像群の統合的変換処理に基づく画像生成の検討", 映像メディア処理シンポジウム (IMPS 2003), I-4.03, pp.73-74, 2003.
- [13] K.Kodama, H.Mo, and A.Kubota: "All-in-Focus Image Generation by Merging Multiple Differently Focused Images in Three-Dimensional Frequency Domain," Pacific-Rim Conference on Multimedia 2005, Part I, pp.303-314, 2005.