

多眼カメラを用いたデジタルリフォーカス技術

Digital Refocusing Technology Using Multiple Fixed-Focus Cameras

山本 琢磨 三島 直 森 達矢

■ YAMAMOTO Takuma

■ MISHIMA Nao

■ MORI Tatsuya

近年、携帯端末の薄型化が進んでいるが、内蔵しているカメラの厚みが原因で更なる薄型化が難しくなっている。カメラの厚さを制限している要因の一つがオートフォーカス機能のためのアクチュエータである。

東芝はこれを解決するため、2台以上のカメラ（以下、多眼カメラと呼ぶ）で撮影した画像の焦点を後から信号処理で変えられるデジタルリフォーカス技術を開発した。この技術により、アクチュエータのない固定焦点カメラにオートフォーカス機能を持たせることができ、携帯端末を更に薄型化できる。

The ongoing dissemination of mobile terminals with thinner profiles in recent years has now reached a stage at which it has become difficult to further reduce the thickness of such terminals due to the thickness of their onboard camera. One of the critical issues in achieving thinner cameras is the size of the actuator that controls the autofocus operation in the camera module.

As a solution to this issue, Toshiba has developed a digital refocusing technology using multiple fixed-focus cameras that allows users to select the focal point after an image is captured. This technology makes it possible to add an autofocus function to fixed-focus cameras, and will contribute to the realization of thinner mobile terminals.

1 まえがき

近年、携帯端末の薄型化が進んでいる。しかし、搭載されているカメラモジュールは、オートフォーカス機能を実現するためのアクチュエータが原因の一つとなって薄くするのが困難になっており、携帯端末自体の更なる薄型化が難しくなっている。東芝は、これを解決するため、アクチュエータを排除して焦点を固定したカメラで撮影した画像を後から信号処理でリフォーカスするデジタルリフォーカス技術を開発した。

デジタルリフォーカスの従来方式は大きく分けて三つある。

第1は、複数の画像を用いる方式である。この方式では、様々な位置にフォーカスした画像を撮影し、被写体ごとに合成処理することで、リフォーカス画像が生成できる⁽¹⁾。しかし、画像の正確な領域分割が必要で、被写体が複雑に重なり合ったシーンへの適用が難しいという問題がある。また、時間的に焦点の位置を変えて撮影するため、被写体が動いた場合に特別な処理が必要となる。

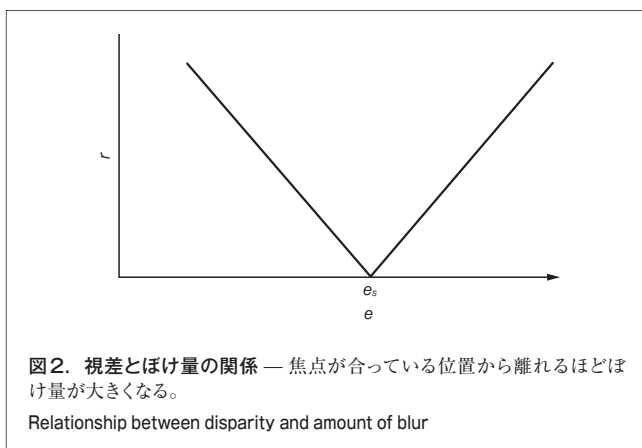
第2は、1枚の画像だけを用いる方式である。画像1枚だけを用いるため、被写体の動きは問題とならない。この方式では、画像の局所領域ごとの高周波数成分の大きさからぼけ具合を推定し、画像をぼかしたり（ぼけの付加）、鮮鋭化したり（ぼけの復元）することでリフォーカス画像が生成できる⁽²⁾。しかし、精細なテクスチャがレンズによってぼけたものなのか、それとも、もともとぼけたテクスチャなのかを区別できないため、単一の画像からはロバストにぼけ具合を推定することは難しい。

第3は、カメラの光学系に特殊な工夫をしたコンピュータショナルカメラ⁽³⁾を用いる方式である。例えば、ライトフィールドカメラ⁽⁴⁾は、センサとメインレンズの間にマイクロレンズアレイを設置することで、昆虫の目で見たような複眼画像を取得し、これを再合成することで、任意の位置に焦点を合わせた画像が生成できる。しかし、カメラ内部を改変するため、製造プロセスの変更が必要となる。

ここでは、焦点を固定した多眼カメラを用いたデジタルリフォーカス技術について述べる⁽⁵⁾。この技術では、一般的な光学系のカメラ2台を平行に並べた通常のステレオカメラを用いるためカメラ内部の改変は必要なく、撮影は1度なので被写体の動きにも影響されない。また、画素ごとにぼけの大きさを推定するため、領域分割の必要もない。類似の手法は、シーンの距離情報（Depth値）が既知であるCG（コンピュータグラフィックス）レンダリングの分野では古くから研究されてきたが⁽⁶⁾、Depth値が未知の自然画に対する従来研究はほとんどない。

2 視差とぼけの大きさの関係

ステレオカメラを用いると、被写体の位置が少しずれた2枚で1組の視差画像が撮影できる。視差画像の例を図1に示す。視差画像1のある画素に対応する点を視差画像2の中から探索により検出すると、画素ごとに被写体のずれ（視差）が求められる。更に、求めた視差から三角測量の原理により被写体までの距離が計算できる。



一般に、カメラの焦点が合っている位置から被写体が離れるほどぼけは大きくなる。レンズの結像公式により距離からぼけの大きさ（以下、ぼけ量と呼ぶ）が計算できる。したがって、視差と距離の関係によって、視差からぼけ量が計算できることになる。視差を e とすると、ぼけ量 r は式(1)で表される。

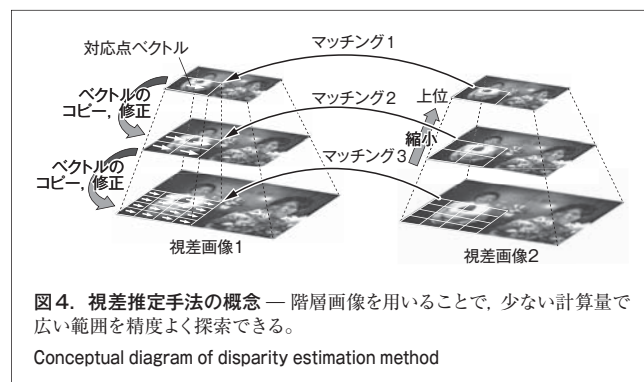
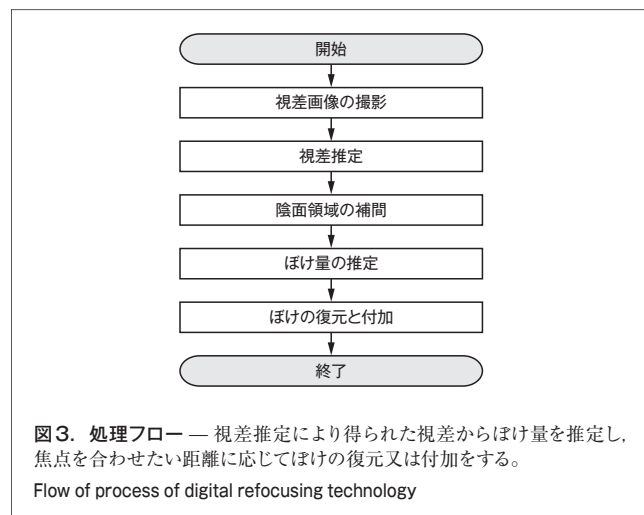
$$r = a|e_s - e| \quad (1)$$

ここで、 e_s はカメラの焦点が合っている距離 s に対応する視差、 a はステレオカメラのカメラ間の距離及び開口の大きさによって決まる定数である。この関係式をグラフで表すと図2のようになる。 e_s から離れるほど、 r が大きくなることがわかる。

3 デジタルリフォーカス技術

当社が開発した方式の処理フローを図3に示す。

まず、ステレオカメラを用いて視差画像を撮影する。次に、視差画像から対応点探索により画素ごとの視差を推定する。更に、陰面領域と呼ばれる、被写体が重なって原理的に対応点が検出できない領域を、その周辺で検出した視差を用いて補間する。そして、式(1)に基づいてぼけ量を推定する。最後に、推定したぼけ量を用いて、焦点を合わせたい距離に応じて画素ごとにぼけの復元又は付加をする。

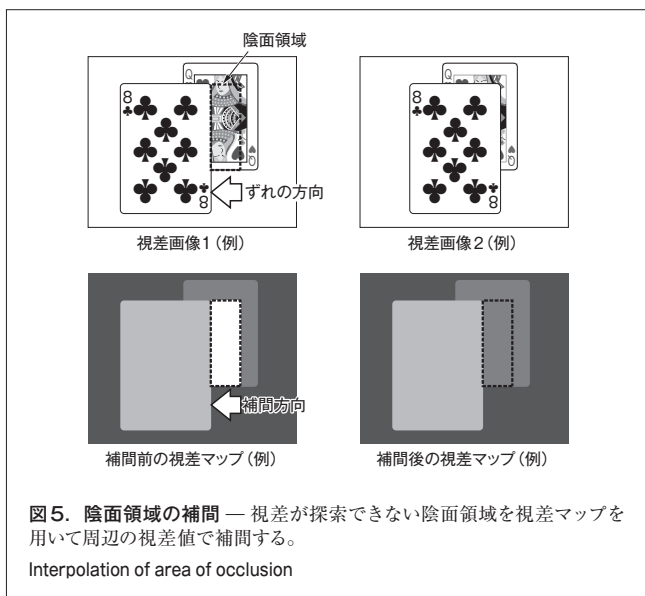


3.1 視差推定

視差は当社が独自に開発した高精度動き推定手法⁷⁾により推定する。まず、図4に示すように、処理対象となる視差画像をそれぞれ段階的に縮小し、複数の縮小画像で構成される階層画像を生成する（図は階層数3の例）。次に、同じ階層の画像どうしで図柄の類似性を基準としたマッチングにより、視差画像1の小領域から視差画像2への対応点ベクトル（小領域内の矢印）を検出する。

ここで、縮小された上位の階層は解像度が低いため少ない計算量で広い範囲を探索できるが、おおまかな対応点ベクトルしか検出できない。一方、下位の階層は解像度が高いため精密な対応点ベクトルが検出できるが、広い範囲を探索しようとすると計算量が多くなる。そこで、上位の階層で検出した対応点ベクトルを一つ下位の階層にコピーし、その対応点ベクトルの指し示す先の周辺だけで探索して、ベクトルを修正する。こうすることで、少ない計算量で広い範囲を精密に探索できる。視差は対応点ベクトルの絶対値（長さ）を計算することで得られる。これにより、視差マップと呼ぶ、視差の大きさを明度で表した画像を作成する。明るいほうが手前になる。

視差画像の中には、陰面領域が存在することがある。例えば前後に並んだ2枚のトランプを撮影した場合、図5上段に



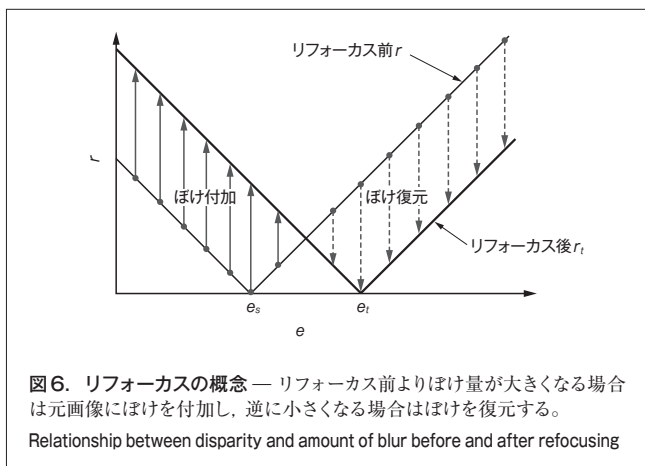
示すように視差画像2には写っていない領域が、視差画像1には存在する。これが陰面領域で、対応点が求まらず視差が検出できない。そこで、陰面領域以外の視差推定をした後に、視差画像間で被写体のずれの方向を特定し、図5下段に示すように、視差マップ上の陰面領域に対して視差データをずれの方向に補間する。

3.2 デジタルリフォーカス

デジタルリフォーカスとはカメラで撮影した画像を s とは異なる、ある指定した距離 t に焦点を合わせた画像に変換することである。 t に対応する視差を e_t とすると、ぼけ量 r_t は式(2)となる。

$$r_t = a|e_t - e| \quad (2)$$

r 及び r_t を図6に示す。これは、リフォーカス前後の視差とぼけ量の関係を表している。画素ごとに、推定したリフォーカス前のぼけ量よりリフォーカス後のぼけ量が大きい場合は画像をぼかすフィルタをかけて(ぼけの付加)、逆にぼけ量が小さ



くなる場合は逆フィルタをかける(ぼけの復元)。画像をぼかすフィルタはレンズのぼけが2次元ガウス関数でモデル化できることから、ガウシアンフィルタを用いる。また、逆フィルタはガウシアンフィルタによるぼけ過程の逆問題を解くことで、設計できる。これにより、リフォーカス画像が生成される。

4 実験

1度の撮影で得た画像から、任意の位置に焦点を合わせた画像を生成する実験の結果について述べる。携帯端末への応用を想定し、5Mピクセルで1/4型のカメラモジュール二つを平行に設置したステレオカメラを用いた。カメラ間の距離は8.5mmとした。また、計算量を削減するために視差推定は、縦横1/2に縮小した画像で行った。

カメラから距離60cm及び140cmに焦点を合わせたリフォーカス画像を図7に示す。また、図7に示した枠内を拡大して図8に示す。焦点を合わせた距離にある物体の鮮鋭感をもっとも高く、その距離から離れるほどぼけており、複雑な

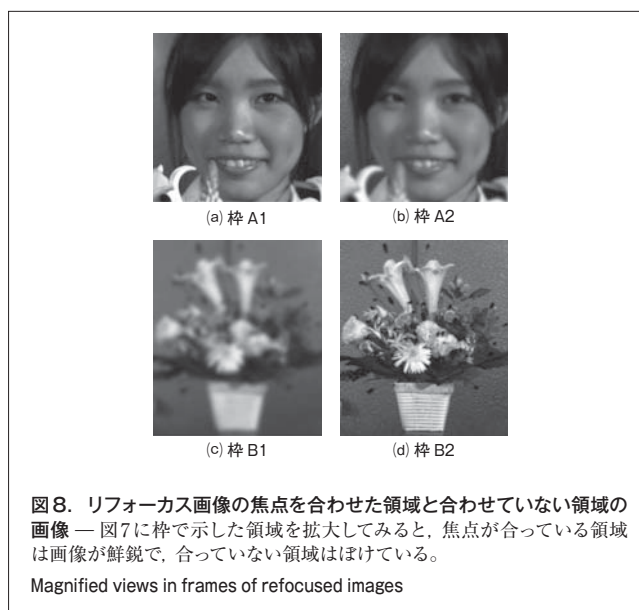
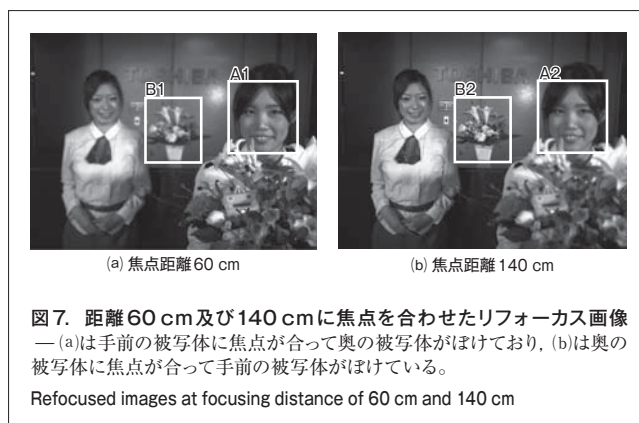




図9. 全焦点画像 — 手前から奥まで全ての被写体に焦点が合っている。
Fully focused image



図10. 検出した視差マップ — 人物や花束の輪郭が明確に検出できている。
Disparity map

シーンであっても期待どおり焦点が調整されていることがわかる。また、一般に携帯端末向けのカメラモジュールは開口が小さいため被写界深度^(注1)が深く、全体に焦点が合ったように見える。これに対し、開発した方式では焦点位置以外でのぼけが強調され、一眼レフカメラで撮影したような印象的な画像が得られ、その有用性が確認できた。

リフォーカス後のぼけ量を視差によらず0に設定することで、全焦点画像も生成できる。生成結果は、図9に示すように、画像全体で鮮鋭感が高く、全ての物体に焦点が合っていた。

視差推定により検出した視差マップを図10に示す。人物や花束の輪郭が明確に検出できており、リフォーカスに用いるのに十分な精度であることがわかる。

5 あとがき

多眼カメラを用いたデジタルリフォーカス技術について述べた。この技術では、視差画像から画素ごとのぼけ量を推定してぼけの付加又は復元をするため、領域を分割する必要がなく複雑なシーンにも適用できる。また、通常のステレオカメラを用いるため、カメラ内部の改変は必要なく、ハードウェア構成がシンプルな利点がある。実験により、複数の被写体が重

(注1) 焦点が合っているように見える被写体までの距離範囲。

なった複雑なシーンであっても任意の位置に焦点を合わせたリフォーカス画像が生成できることを確認した。

この技術により、固定焦点のカメラであってもオートフォーカス機能を付与できるため、携帯端末の更なる薄型化が期待できる。また、視差マップから直接被写体までの距離を計算して明度で表した距離画像が得られるので、被写体を輪郭に沿って切り出すオブジェクト抽出や、ジェスチャ入力などへの応用も期待できる。

文献

- (1) Hasinoff, S. W. ; Kutulakos, K. N. "A Layer-Based Restoration Framework for Variable-Aperture Photography". Proc. 11th Int. Conf. on Computer Vision (ICCV). Rio de Janeiro, Brazil, 2007-10, IEEE. 2007, p.1-8.
- (2) Zhang, W. ; Cham, W. K. "Single Image Focus Editing". 12th. ICCV Workshops. Kyoto, Japan, 2009-09, IEEE. 2009, p.1947-1954.
- (3) Levin, A. et al. Image and Depth from a Conventional Camera with a Coded Aperture. SIGGRAPH. 26, 3, 2007, p.70-1-70-9.
- (4) Ng, R. et al. Light field photography with a hand-held plenoptic camera. Stanford U. Tech. Rep. Computer Science Technical Report (CSTR). 2005-02, p.1-11.
- (5) 山本琢磨 他. "ステレオカメラを用いたデジタルリフォーカス". Vision Engineering Workshop (ViEW) 2013. 横浜, 2013-12, 精密工学会画像応用技術専門委員会. 2013, IS1-F1.
- (6) Barsky, B. A. ; Kosloff, T. "Algorithms for rendering depth of field effects in computer graphics". 12th World Scientific and Engineering Academy and Society (WSEAS) International Conference on Computers. Heraklion, Greece, 2008-07, WSEAS. 2008, p.999-1010.
- (7) Yamamoto, T. et al. "High-Accuracy Motion Estimation with 4-D Recursive Search Block Matching". 1st IEEE Global Conference on Consumer Electronics (GCCE). Tokyo, Japan, 2012-10, IEEE. 2012, p.625-628.



山本 琢磨 YAMAMOTO Takuma

研究開発センター マルチメディアラボラトリー。
映像システムやイメージセンサ向けの画像処理アルゴリズムの研究・開発に従事。電子情報通信学会会員。
Multimedia Lab.



三島 直 MISHIMA Nao

研究開発センター マルチメディアラボラトリー主任研究員。
画像処理技術の開発に従事。
Multimedia Lab.



森 達矢 MORI Tatsuya

セミコンダクター&ストレージ社 半導体研究開発センター
デジタルメディアSoC技術開発部主務。画像処理及び画像認識の研究・開発に従事。
Center for Semiconductor Research & Development