

高密度指向性表示 3D ディスプレイにおける簡易実写入力手法

壺井 雅史 木村 真治 堀越 力

株式会社 NTTドコモ 総合研究所 〒239-8536 神奈川県横須賀市光の丘 3-5
E-mail: {tsuboim, kimurashi, horikoshit}@nttdocomo.co.jp

概要 高密度指向性表示 3D ディスプレイにおける実写入力手法を提案する。提案手法では、静止している物体を撮影対象とし、単独のカメラで撮影した動画から自然特徴点を抽出することで動画中に含まれる異なる視点画像同士を対応付け、立体映像を生成する。本手法は、得られた特徴点同士の対応関係から、簡易的な幾何制約のみを用いて立体映像を生成する。そのため、幾何学的制約を全て満たすわけではないが、高密度指向性表示方式の 3D ディスプレイに立体映像を表示するための素材としては十分な映像を生成でき、かつ簡易な操作のみで立体映像を取得することができる。

キーワード 3D ディスプレイ, 実写入力, 高密度指向性表示

A method to input a *real* scene for a 3D display using high-density directional images

Masashi TSUBOI Shinji KIMURA Tsutomu HORIKOSHI

NTT DoCoMo Inc. Research Labs. 3-5 Hikarino-oka, Yokosuka-shi, Kanagawa, 239-8536 Japan
E-mail: {tsuboim, kimurashi, horikoshit}@nttdocomo.co.jp

Abstract We propose a method to input a real scene for a 3D display using high-density directional images. The 3D image is created by using a correspondence between the frame images of the movie of the static object. Our method uses some parts of the geometric constraints; therefore the correct image cannot be generated. However, 3D image created by our method has sufficient quality for a 3D display using high-density directional images. Moreover, 3D image can be created only by easy operations, which are feature detection and translation of the frame images, because our method does not need any complex computation.

Keyword 3D display, Real scene input, high-density directional images

1. はじめに

近年、自然で臨場感のある立体映像を提示できる 3D ディスプレイの実現に対する期待が高まっている。特に我々は、モバイル用途の 3D ディスプレイとして、高密度指向性表示を用いたプロトタイプを試作し、カメラや傾きセンサを用いて上下左右に自然な運動視差を実現する手法を提案した[1]。

将来の立体映像コミュニケーションにおいては、現実の立体情報を簡易に取得できること

が重要である。特に我々の試作した 3D ディスプレイで採用している高密度指向性表示方式の場合、実写立体映像を表示するためには原理的に数多くのカメラを必要とし、さらに画像補間などの処理が必要となる。

本稿では、立体映像を高密度指向性表示方式の 3D ディスプレイに表示するための簡易的な実写映像入力方式を提案する。提案手法では、静止している物体を撮影対象とし、単独のカメラで撮影した動画から自然特徴点を抽出するこ

とで動画中に含まれる視点画像同士を対応付け、立体映像を生成する。本手法は、得られた特徴点同士の対応関係から、簡易的な幾何制約のみを用いて立体映像を生成するため、幾何学的制約を全て満たすわけではないが、高密度指向性表示方式の 3D ディスプレイに立体映像を表示するための素材としては十分な品質の映像を生成でき、かつ複雑な計算を含まないので高速に立体映像を取得することができる。

2. 指向性画像表示方式

高密度指向性表示方式では、図 1 に示すように複数の方向(#1-#5)に準平行光として画像が表示される(この各画像を指向性画像と呼ぶ)。指向性画像を表示する方向を非常に細かくすることで、微小な観察位置の移動であっても異なる指向性画像が観察されることとなり、滑らかな運動視差を実現できる。特に、通常の液晶パネルに、斜めにレンチキュラーシートを貼り付けた薄型の構成によって、水平方向に高密度指向性表示が実現できる [2]。

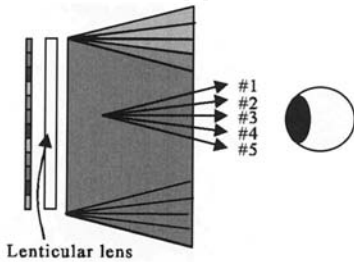


Figure 1: High-density directional images

高密度指向性表示方式 3D ディスプレイで表示される立体映像は、観察者の網膜上では、ディスプレイ表面を短冊状に区切った領域ごとに

異なる指向性画像組み合わせたものとなる。すなわち、観察者は網膜上でひとつの指向性画像のみを観察するのではなく、指向性画像の組み合わせとして画像を観察することになる(図 3)。指向性画像表示刻みは、この短冊状の領域の幅に影響し、表示刻みが細かければ、短冊状領域の幅も細くなるために、観察者の位置の変化に伴いより滑らかな運動視差を実現できる。

我々は、市販の液晶ディスプレイと斜めレンチキュラーシートを組み合わせ、3種のモバイル型 3D ディスプレイを試作した。試作した 3種のディスプレイの仕様を以下の表 1 に示す。

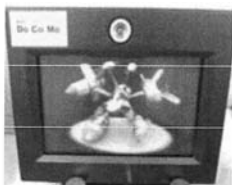
我々の試作した 3D ディスプレイは、いずれも異なる方向の指向性画像を約 1 度間隔で表示するため、滑らかな運動視差を実現することができる。

3. 目的

我々の目的は、上述した高密度指向性表示 3D ディスプレイにおいて、自然な実写映像を入力することにある。そのためには、多数の指向性画像が必要となるが、多数のカメラを用いる手法は現実的ではない。そこで、単独のカメラから撮影した動画をもとに指向性画像を取得することを考える。カメラを水平に動かして動画を撮影し、各フレームを切り出す。そうすることによって、水平方向に平行に並べた複数のカメラにより得られた多視点画像とほぼ同等の画像を得ることができる。しかし、実際には、ユーザが手で持ってカメラを動かす場合、カメラを理想的に平行に移動させることは不可能であり、カメラの位置ズレや、光軸のブレが発生する(図 4)。

Table 1. Specifications of the prototyped displays

	パネルサイズ	パネル解像度	指向性画像数	指向性画像表示刻み
1	7.2 [inch]	1280 x 768	30	0.94 [deg]
2	4.5 [inch]	1024 x 600	30	1.16 [deg]
3	2.5 [inch]	480 x 690	24	1.54 [deg]



(a) type 1 (7.2inch)



(b) type 2 (4.5inch)



(c) type 3 (2.5inch)

Figure 2. Prototyped 3D displays

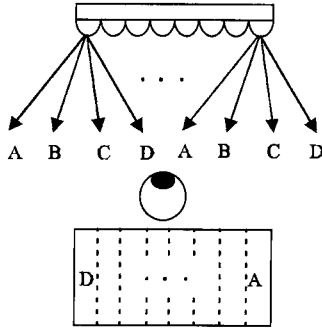


Figure 3. Retinal image

本研究の目的は、こういったブレの存在を前提として、高密度指向性表示方式の3Dディスプレイに適した指向性画像を得ることにある。ここで、いつでもどこでも利用可能なモバイル環境を想定して、撮影対象物は未知のものとし(すなわち、既知のパターンを利用しない)、カメラは現在のモバイル機器に搭載される程度の性能(QVGA~VGA解像度)を持つものとする。

4. 既存手法

水平方向にカメラを動かして撮影対象物を撮影した場合、動画から切り出された各フレームは、未校正であり、かつ位置、光軸の変動は大きい。ほぼ水平に並んだカメラアレイにより撮影された画像であるといえる。

このように、異なる位置にあるカメラ画像間の対応関係を求め、画像を補正するという要求は古くから検討されている。例えば、空間中の光線情報を記述し、実写映像から仮想視点の画像を生成する手法である Light Field Rendering[3]や、多数の実写映像を組み合わせるパノラマモザイク[4]などの、実写映像を取り扱う基本的な手法として検討されてきた。また、実写画像同士の対応付けを行う手法として、エピポーラ幾何に基づく画像の平行化(rectification)を行う手法[5]がある。しかし、これらの手法を本研究の目的となる手ブレなどを含む自然画像から指向性画像を得る手法に適用するのは難しい。

例えば、Dengらは2次元平面上を自由に動くカメラを前提とし、位置ブレを前提とした画像の平行化を行っている[6]が、機械的にカメラを平行移動させるため、カメラの姿勢は原則として変わらない上、同心円状の既知パターンによ

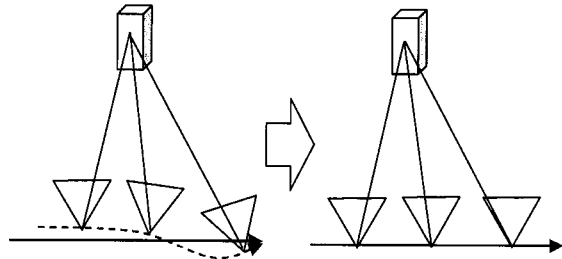


Figure 4. Correction of captured images

り射影行列を推定するため、自然画像のみから画像を補正するのは難しい。また、パノラマモザイクのために、カメラからみて無限遠に存在する平面を利用してカメラの内部パラメータと姿勢を均一にそろえる手法も提案されている[4]が、必ずしも無限遠に存在する平面を抽出できるわけではない。特に本研究のように立体表示を目的とした場合、風景のように無限遠平面を含むことが期待できる画像よりも、比較的手元に近い位置に存在する物体を撮影することのほうが多い。

このように、既存の手法を用いて本研究の目的である立体映像表示に適した指向性画像を得ることは容易ではない。未校正カメラを前提とし、複数の画像上からエピポーラ幾何をもとにしてF行列(Fundamental Matrix)を求め、画像間の対応関係を求める手法もある[7]が、特徴点の精度などに非常に大きく影響を受けるデリケートな計算を要求される。そこで、本研究では、最大の目的は観察者にとって違和感の少ない立体映像を提示することである、という目的に沿って、簡略化した制約条件のもとで高速に立体映像を得る手法を提案する。

5. 提案手法

本提案手法は以下のステップからなる(図5)。
ステップ1: 撮影動画中から各フレームの画像を切り出し、画像に含まれる自然特徴点を抽出する

ステップ2: ステップ1で求めた自然特徴点のうち、全画像に共通の特徴点を求める

ステップ3: ステップ2でもとめた共通の特徴点を基準として各画像の位置を調整する

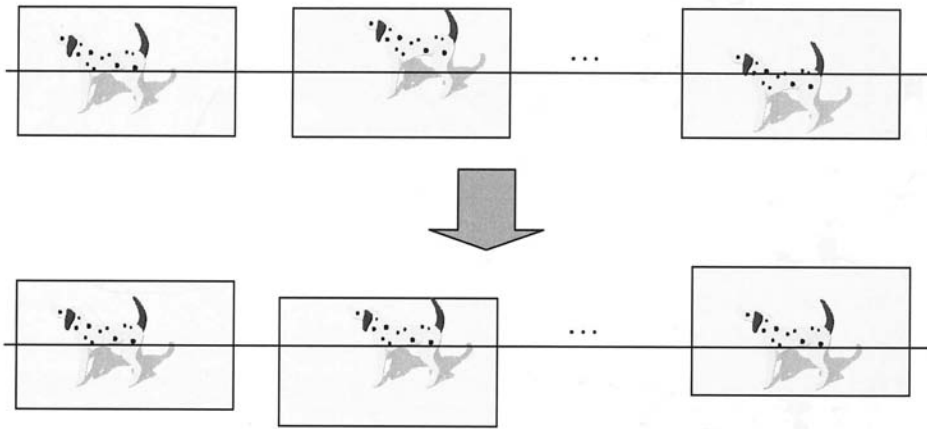


Figure 5. Algorithm

ステップ1 およびステップ2 では、動画に含まれる各画像に共通の特徴点を抽出する。ここでは、自然特徴点の抽出には SIFT (Scale Invariant Feature Transform) を用いた[8]。SIFT は、画像中の特徴点を 128 次元のベクトルで表現しており、画像の回転や拡大縮小に対してロバストに特徴点を抽出できるという特徴を持つ。

画像 i 中の画素 (x_i, y_i) における SIFT 特徴量ベクトルを $\vec{s}_i(x_i, y_i)$ とおくと、別の画像 j 中の画素 (x_j, y_j) との特徴量の差は

$$\Delta s = \sqrt{|\vec{s}_i(x_i, y_i) - \vec{s}_j(x_j, y_j)|^2}$$

として求められる。求めた Δs が一定量以下である場合、同一の特徴点であるとみなす。ステップ2 においては、全ての画像で共通の特徴点を求める。前記の条件のみでは共通の特徴点が存在しなかった場合には、特徴点对応のとれなかった画像については対応のとれた点同士の情報から補間して仮想的に特徴点の位置を求めることとする。

ステップ3 では、ステップ2 で求めた共通特徴点がすべて同じ位置にくるように画像を補正する。当然、実際の映像を用いた場合にはオクルージョンなどが発生し、画像の平行移動のみで正確な画像を得ることは不可能である。しか

し、2 節で示したとおり、高密度指向性表示 3D ディスプレイにおいては、観察者は網膜上で個別の指向性画像を観察するのではなく、複数の指向性画像を組み合わせるため、擬似的に生成した画像で、手ブレなどに起因する違和感を抑えた立体感を与えることは可能である。

6. 実験

提案手法を用いて、立体映像の生成を行った。動画の取得には、NEC 製携帯電話 N903i に搭載されているカメラを用い、VGA 解像度で約 3 秒間の動画を取得し、その中の 30 フレームを抽出した。動画から抽出されたフレーム画像の一部を図に示す。これらの画像に対して SIFT 特徴点を抽出した。各画像から平均 57 点の特徴点が抽出され、このうち、全画像に共通の特徴点は、図 6(a) に示すように、5 点あった。また、他のシーンを撮影した動画に対しても、共通の特徴点が数点程度抽出できることを確認できた。

これらの共通特徴点のうち、一点を選択し、全てを同じ位置となるよう補正した。補正済み画像を図 6(b) に、立体映像表示した模様をカメラで撮影した画像を図 7 に示す。

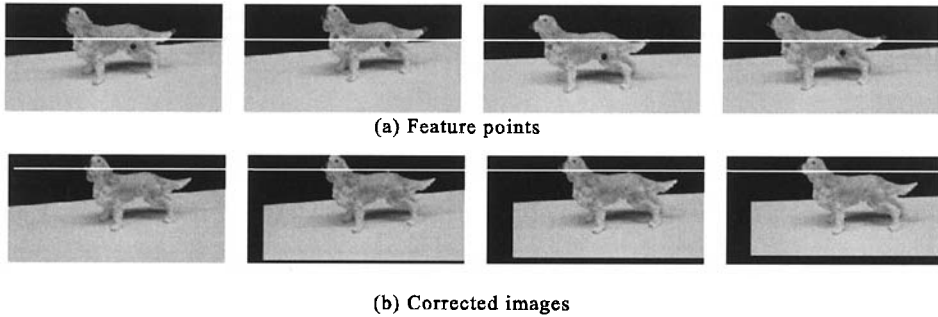


Figure 6. Feature points and corrected images

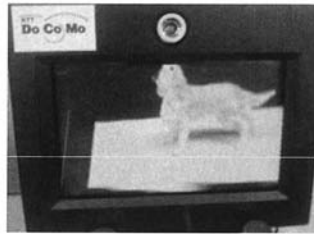


Figure 7. Represented 3D image

7. 考察

提案手法により生成された指向性画像は、原理的に幾何学的拘束を厳密に満たすわけではない。しかし、高密度指向性表示 3D ディスプレイでは、観察者は個別の指向性画像を観察するのではなく、複数の指向性画像を短冊状に組み合わせた画像を観察する。また、本研究の目的は、複雑な計算を行うことなく、立体表示に適した指向性画像を得ることである。このため、提案手法の有効性を示すためには、生成した立体映像をこの目的に従って評価する必要がある。

例えば、高密度指向性表示を用いた 3D ディスプレイにおいては、非常に細かく指向性画像を表示することにより、金属の光沢や果実のみずみずしさを表現することができる。この点に着目し、立体物の質感表現能力に着目した主観評価[9]が行われている。そのほかにも、観察者の目の調節応答に着目し、立体映像の奥行認識とあわせた人間工学的な評価が行われている[10]。このように、立体映像を評価するためには従来の SNR ではなく、こういった各方式特有の特徴に基づいた評価方法および評価項目を検討する必要がある。

特に、提案手法は、観察者が指向性画像を短冊状に切り取った組み合わせの画像を観察するという点に着目して、考慮すべき幾何学的条件を簡略化している。そのため、提案手法を評価するにあたっては、この簡略化の妥当性を評価すべきである。すなわち、個々の指向性画像を単体で評価するのではなく、観察者の網膜像を客観的に評価する必要がある、また、立体映像表示の結果としてあらわれる立体感や違和感を主観的に評価していく必要もあると考えられる。

8. まとめと今後の課題

高密度指向性表示を用いた 3D ディスプレイに表示するための、簡易的な実写立体映像入力手法を提案した。提案手法は、単眼カメラにより撮影された動画をもとに、自然特徴点を利用することで必要な指向性画像を生成するため、既知のパターンなどを取得する必要がなく、簡易な操作で立体映像を生成することができる。

今後、主観評価により、生成した実写立体映像の評価を行い、指向性画像表示 3D ディスプレイにおける実写映像入力における、本手法の有性を検討していく予定である。

謝辞

本研究において、3D ディスプレイの開発にご協力いただいた東京農工大学 高木康博准教授に感謝いたします。

文 献

- [1] 壺井雅史, 藤岡正樹, 高木康博, 堀越力: “モバイル用高密度指向性三次元ディスプレイの視域拡大と擬似フルパララックス化”, 3次元画像コンファレンス 2006, pp.53-56, (2006)
- [2] Y. Takaki : “Thin-type natural three-dimensional display with 72 directional images”, Proc. of SPIE, vol.5664, pp.56-63, (2005)
- [3] M. Levoy, et al.: “Light Field Rendering”, Proc. ACM SIGGRAPH’96, pp.31-42, (1996)
- [4] R. Szeliski : “Video Mosaics for Virtual Environments”, IEEE Computer Graphics and applications, March 1996, pp. 22-30, (1996)
- [5] N. Ayache, C. Hansen : “Rectification of images for binocular and trinocular stereo vision”, Proc. 9th Int. Conf. Pattern Recog., pp.11-16, (1988)
- [6] K. Deng, L. Wang, Z. Lin, T. Feng, and Z. Deng: “Correction and rectification of light fields”, Computers & Graphics., 27, pp.169-177(2003)
- [7] 菅谷保之, 金澤靖, 金谷健一: “エッジ極線幾何学による 2 画像間の密な点対応の生成”, 情報処理学会研究報告 2005-CVIM-148-19, pp.145-152, (2005)
- [8] David G. Lowe: “Object recognition from local scale-invariant features”, Proc. Int. Conf. on Computer Vision, pp.1150-1157, (1999)
- [9] 大力孟司, 畑田豊彦, 高木康弘: “VGA 解像度高密度指向性ディスプレイによる質感表現に対する主観評価”, 3次元画像コンファレンス 2006, pp.69-72, (2006)
- [10] 村瀬祐二, 栗林英範, 大力孟司, 高木康弘, 畑田豊彦: “視差情報量の異なる立体表示方式での空間再現状態の評価”, 映像情報メディア学会技術報告 Vol.31, No.18, pp.1-4, (2007)