

高臨場感仮想体験のための装着型全方位撮影システム

Wearable Imaging System for Capturing Omnidirectional Movies from First-person Perspective

近藤 一晃 (京都大学), 向川 康博 (大阪大学), 八木 康史 (大阪大学)

Kazuaki Kondo(Kyoto Univ.), Yasuhiro Mukaigawa(Osaka Univ.), and Yasuhi Yasi(Osaka Univ.)

Abstract We propose a novel wearable imaging system that can capture omnidirectional movies from the viewpoint of the camera wearer. The imaging system solves the problems of resolution uniformity and gaze matching that conventional approaches do not consider. We combine cameras with curved mirrors that control the projection of the imaging system to produce uniform resolution. Use of the mirrors also enables the viewpoint to be moved closer to that of the camera wearer, thus reducing gaze mismatching. The optics, including the curved mirror, have been designed to form an objective projection. The capability of the designed optics is evaluated with respect to resolution, aberration, and gaze matching. We have developed a prototype based on the designed optics to capture various scenes from first-person perspective.

1 はじめに

我々が日常生活で目にする映像や画像は、ほとんどの場合シーンに対して客観的な第三者視点から撮影されたものである。このようなカメラワークは視聴者がストーリーやシーンの構成を理解するのに役立つ。それに対しシーンに登場する人物の視点は主観視点と呼ばれ、臨場感を与えたい場合の撮影に適している。さらに、迫力のある臨場感を与える手法としては自己の周囲 360° を覆う全方位映像の利用も考えられる。それら 2 つを組み合わせた主観視点全方位映像を利用することで、視聴者にまるでその場にいるかのような感覚を与えることができる。そのため主観視点全方位映像を用いた仮想体験システムがこれまでにいくつか提案されてきた。山澤ら [1] は人物の頭部に全方位カメラを搭載することで、人物視点の全方位映像を獲得している。また尾上ら [2] や池田 [3] らは車両やジェットコースターといった動物体を用いて全方位撮影を行い、視聴者の四方を囲む全方位呈示系とともにテレプレゼンスシステムを開発している。

主観視点全方位映像を用いて臨場感の高い仮想体験を行うためには、以下に挙げる 2 項目の達成が重要である。第一に全方位映像に均一な分解能を持たせること、第二に撮影者の視線と撮影系の視線を一致させることである。全方位映像が呈示されている状況では視聴者は見る方向を自由に変えることができる。任意の方向に対して一定画質の映像を提供するためには、均一な分解能を持つ全方位映像が求められる。また撮影者と共演者が向かい合って会話しているシーンなどでは、目の動きやアイコンタクトといった視線の動きが大きな役割を持つ。視線を正確に表現するためには、まさに撮影者の視点か

ら見た映像が要求される。

これまでに提案された全方位カメラは主観視点からの撮影を想定していないため、ほとんどの場合上記項目を満たしていない。凸型反射鏡を用いた反射屈折光学系は方位方向には均一な分解能を持つが、縦方向に関して不均一である。またカメラを 1 台しか使っていないため全体の解像度が不十分である。マルチカメラで構成された全方位カメラは十分な解像度を提供できるが、分解能の均一さを考慮してはいない。さらに両者ともに人物への装着を想定していないため、装着者の視点に近づけることが困難で十分な視線の一致が得られない。加えて映像の記録には別途記録用の装置を必要とする。映像や電源を有線で伝達することも可能であるが、撮影の自由度を著しく阻害してしまうため実用的でない。

そこで我々はこれらの問題を解決した主観視点全方位映像撮影用の新たな装着型システムを提案する。すなわち提案システムは撮影者とほぼ同じ視点から見た均一な解像度を持つ全方位方位映像の獲得を目指す。また全体的な解像度や視野といった全方位画像本来の特性や、自由度の高い撮影のためのウェアラブル性能も同時に満たす。撮影システム全体としての仕様を明らかにした上で、それらを達成するように設計を行い、試作を通して映像の評価を行う。

2 目標仕様設定

目的の撮影システムが満たすべき仕様は表 1 に示すとおりで、光学的仕様と機能的仕様に分類される。

表 1: 提案システムの目標値と従来手法によるその達成度．マルチカメラ系として Ladybug2[6]，反射屈折光学系として HyperOmni-Vision($a = b = 20mm$)[5] を従来手法の代表例として挙げている．最右列は提案手法に基づいて開発した試作システムのスペックを示す．

terms	requirements	Ladybug2	HyperOmni-Vision	our system
Horizontal FOV	360°	360°	360°	360°
Vertical FOV	-30° ~ +25°	-36.9° ~ +36.9°	-90° ~ +25.0°	-31.8° ~ +26.6°
Resolution(horizontal)	3840pixels	3580pixels	0 ~ 3220pixels	3960 ~ 5024pixels
Resolution(vertical)	590pixels	768pixels	520pixels	584 ~ 682pixels
Resolution(distribution)	uniform	did not consider	did not consider	uniform
gap of viewpoints	4.4cm	over 18cm	over 15cm	5.0cm
Size	W60cm × D40cm	W10cm × D11cm	W10cm × D10cm	W40cm × D40cm
Additional equipments	no(stand-alone)	power and recorder	power and recorder	no(stand-alone)
Max. driving time	over 30minutes	with power line	with power line	85min.
Weight	2.0kg	1.2kg	0.5kg	2.4kg

2.1 光学的仕様

(1) 視野の大きさと全体解像度

人間工学の観点から，視野の大きさと全体解像度の目標値を決定する．我々は注視安定視野と呼ばれている視野を選択した．これは作業をする場合やシーンを鑑賞する場合に形成される自然な視野で，横，上，下の方向へ各々， $\pm 30^\circ \sim 45^\circ$ ， $20^\circ \sim 30^\circ$ ， $25^\circ \sim 40^\circ$ である [7]．全方位映像では横方向の視野は当然 360° であるため，上下に関する視野角の目標値を上記を元に決定した．全体解像度は横方向への視野角 (60°) に，コンピュータビジョンで一般的に使用されている VGA 画像の横方向解像度 ($640pixel$) を対応させることで決定した．

(2) 分解能の均一さ

視聴者に可視情報を均一に与えるためには，全方位映像における分解能はできるだけ均一に分布していなければならない．映像は二次元の情報なので，縦横の両方向に沿って分解能が均一となることを目指す．全方位呈示系の代表である CAVE システム [4] は平面スクリーンの組合わせで構成されているため，シームレスな映像呈示には不向きである．本論文では映像が滑らかに繋がることを重視し，方位角 (θ) と上下角の正接 ($\tan \phi$) で張られるパノラマシーンを対象として分解能を扱う．

(3) 視線一致度

視線一致度は特に撮影システムの視点と撮影者の視点の縦間隔に深く関係し，一般に隔たりが小さいほど視線が一致しやすい．図 1 は撮影者と共演者が向かい合って相手の目を注視した場合の模式図である．撮影系と撮影

者の視点間距離 H_v が大きいと，共演者の視線を正しく撮影できないため，撮影系の視点はなるべく下げる必要がある．共演者に対して，撮影者ではなく意図して撮影系を見るように演技してもらうことも可能だが，やはり不自然な演技となってしまう．また演技によって視線合致の問題を解決したとしても，撮影者の視点から外れた撮影であることは変わりなく，これでは有効な主観視点映像は得られない．そのため物理的に 2 視点を近づけることが重要である．

我々は隔たり H_v の許容値を以下のようにして決定し，視線一致度の指標とした．隔たり H_v は撮影システムと撮影者の視線の間に視線変位角 $\delta\phi$ を形成する．

$$\delta\phi = \arctan \frac{H_v}{D} \quad (1)$$

心理学的な検証から人間が違和感を感じる視線変位角はおよそ 5.0° と言われている [8]．さらに，同じ机に座って向かい合った状態で会話をするなどの日常における最も近いシーンを想定し，視線を一致させたい奥行き距離を $50cm$ 以遠とした．これらの数値と (1) 式に基づいて視点間の隔たりの許容値を計算するとおよそ $4.4cm$ となる．我々はこの値を視線一致度に関する目標値として採用した．

2.2 機能的仕様

主に装着型の撮影システムを構成するための条件として，大きさは一般的な人間の大きさを元に幅 $60cm$ ，奥行き $40cm$ 以下とした．また記録装置や電源供給を含んだ高い独立性も求められる．有線による電源供給は撮影の自由度を損なうため，バッテリー駆動が望ましい．そ

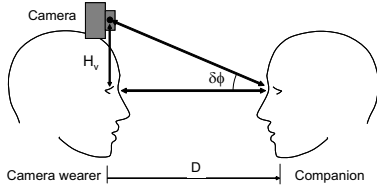


図 1: 撮影者と撮影システムの視点が離れていることによって生じる視線の不一致．図は頭部搭載カメラの場合を表している．

の場合は駆動時間も考慮されるべきで，連続稼働時間の目標値を 30 分とした．重量は一般的に頭部に装着する最重の物体よりも軽くする必要がある．そこでフルフェイスタイプのヘルメットを元におよそ 2.0kg とした．

3 光学系の設計

3.1 基本方針

目的の撮影システムを，記録装置を内蔵した小型カメラと曲面ミラーからなる光学ユニットの組み合わせで実現する．曲面ミラーはカメラの視野角を広げるだけでなく，生成される射影を制御できる．またミラーを用いることで光学系の仮想視点を撮影者の視点位置に近づけることもできる．全体の解像度はカメラの台数を増やすことで簡単に達成できるが，同時に全体の大きさや重量の問題が発生する．そのトレードオフを考慮し 4 台の光学ユニットを用いて撮影システムを構成することにした．基本構成は図 2(a) に示すとおりである．

3.2 ミラー曲面設計法

最適な曲面ミラー形状を，5 つのステップを繰り返すことで設計する．各々の詳細を以下に示す．

(1) 法線マップの算出

まず曲面ミラーが満たすべき理想法線マップをカメラパラメータと目的の視野角から算出する．次に算出された法線マップを形成するような曲面ミラーを推定する．カメラパラメータは，焦点距離 f ，およびカメラの位置姿勢を決定する行列 $\mathbf{P} = [\mathbf{R} \ \mathbf{t}]$ である．図 2(b) は光学ユニットにおけるカメラ，ミラー，シーンの関係を示している．理想法線マップを算出するために，まず目的射影をもとにカメラ像面上の各画素とシーン中の光線との対応を計算する．パノラマシーンにおいて均一な分解能を実現させる目的射影 \mathbf{V}_s は世界座標系において以下の式で与えられる．

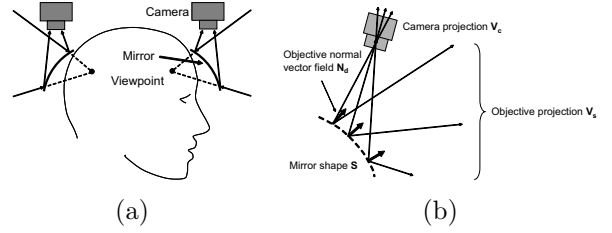


図 2: 提案システムの構成．(a) 基本概念図．システム全体を構成する 4 台の光学ユニットのうちの前 2 台のみを図示している．(b) 光学ユニットの光路図．

$$\mathbf{V}_s(u, v) = \begin{bmatrix} \tan(\frac{u}{U}(\theta_{max} - \theta_{min}) + \theta_{min}) \\ \frac{v}{V}(h_{max} - h_{min}) + h_{min} \\ 1 \end{bmatrix} \quad (2)$$

ここで (u, v) は $U \times V$ の大きさのカメラ像面上における画素の位置を示す．目的視野の大きさは方位方向角度 $\theta_{min}, \theta_{max}$ ，および縦方向角度の正接 h_{min}, h_{max} によってパノラマシーン上の長方形で与えられる．一方，カメラ自身の射影 \mathbf{V}_c は透視投影に基づいて $\mathbf{V}_c(u, v) = [u \ v \ -f]^t$ と表される．これらの射影 \mathbf{V}_s と \mathbf{V}_c は曲面ミラーの反射により関係付けられる．反射の性質から反射点における曲面ミラーの法線は \mathbf{V}_s と \mathbf{V}_c がなす角を二等分する．以上より理想法線マップはベクトル正規化演算子 $\mathcal{N}[\mathbf{x}] = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ を用いて次式で表される．

$$\mathbf{N}_d = \mathcal{N}[\mathcal{N}[\mathbf{V}_s] + \mathcal{N}[\mathbf{R}\mathbf{V}_c]] \quad (3)$$

(2) 曲面ミラーの形状設計

引き続き \mathbf{N}_d を満たすような曲面ミラーの形状を推定する．実際には目的射影を完全に満たすミラー形状は存在しないため最適解を求めることとなる．我々は Swaminathan らによる線形アルゴリズム [9] を用いて最適な曲面ミラー形状を計算した．曲面ミラー形状 $\mathbf{S}(u, v)$ は，2 次元の像面上で定義される 4 次スプライン曲線の積で以下のように表される．

$$\mathbf{S}(u, v) = \mathbf{R}\mathbf{V}_c(u, v) \sum_{i,j} C_{ij} f_i(u) g_j(v) + \mathbf{t} \quad (4)$$

ここで C_{ij} はスプライン曲線の制御点， $f_i(u), g_j(v)$ は 3 次スプライン曲線の u, v 方向の基底関数を示す．曲面上の接線はその点における法線と直交することから

$$\frac{\partial \mathbf{S}}{\partial u} \cdot \mathbf{N}_d = \frac{\partial \mathbf{S}}{\partial v} \cdot \mathbf{N}_d = 0 \quad (5)$$

なる方程式を得る．曲面ミラー上の各点における (5) 式を並べて線型方程式とし，擬似逆行列を用いて解くことで曲面ミラー形状の最適解を最小二乗近似で算出する．

(3) 射影の確認

前項で述べたように，設計された曲面ミラーは目的射影を完全に再現するとは限らず，近似的な射影を形成する場合がほとんどである．そのため，目標射影をどの程度近似しているのかを確認する必要がある．十分な近似が得られていたならば次のステップに進むが，そうでない場合は (5) に示すカメラパラメータの調整を行い，法線マップの算出に戻る．

(4) 収差解析

反射屈折光学系の場合，曲面ミラーの形状は射影だけでなく集光にも影響を与える [10]．そのため設計された曲面ミラーにおける収差を解析し，光学系が適切な集光状態にあるかどうか確認すべきである．デジタル撮像素子を用いたカメラの場合，収差が 1 素子の対角長さ以下に収まっているかどうかで集光状態が決定される．すなわち対象物のスポットダイアグラムの大きさが，対角長さ以下かどうかを調べればよい．収差が大きく十分な集光状態が得られていない場合は，収差量を低減するように調整を行い，法線マップ計算に戻る．

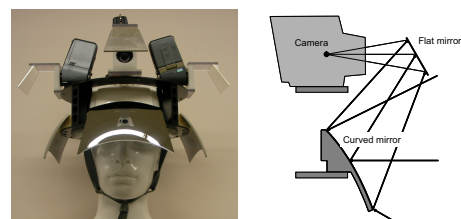
(5) カメラパラメータの調整

本論文で扱う光学構成では，広角カメラを用いることで目的射影に対する近似精度を上げることができる．すなわち f を小さくすればよい．一方，収差は曲面ミラーを大きくすることで低減される．そのためには t を長くすればよい．また R を調整してカメラの縦方向配置角度を水平に近づければ，射影の近似精度，集光の両者を高めるとができる．カメラ筐体によるセルフオクルージョンや光学系の大きさに関する制限などに留意して経験的にパラメータ調節を行った．

4 主観視点全方位撮影システムの試作

4.1 光学系の構成

光学ユニットが満たす目的射影として方位方向に 110° ，縦方向に $\pm 30^\circ$ のパノラマ視野における均一な射影を与え，曲面ミラーの設計を行った．方位方向の目標視野角として， $360^\circ/4 = 90^\circ$ よりも大きく設定したのは，隣接する光学系の視野に重なりを作るためである．重なった視野内の対象物は隣接するどちらの画像にも現れるた



(a) overview (b) optical unit

図 3: 試作機の概観と光学ユニットの構成．

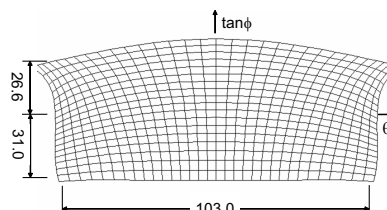


図 4: 設計した光学ユニットにより形成される射影．カメラ像面上の格子パターンをパノラマシーン上に投影した様子を図示している．横軸は方位角 θ ，縦軸は仰俯角の正接 $\tan \phi$ である．

め，画像の結合に利用できる．曲面ミラーの形状表現に用いた制御点の数は縦横 10×8 個である．ミラーの大きさは，およそ $158\text{mm} \times 100\text{mm} \times 25\text{mm}$ となった．

曲面ミラー・カメラ・追加の平面ミラーを組み合わせて光学ユニットを構成し，それを前後左右の放射状に配置することで試作機を作成した (図 3)．平面ミラーは全体をコンパクトに収めるために用いたが，集光や射影といった光学的な性質には影響を与えない．我々はこの試作機を主観視点全方位撮影システム First-Person Perspective Omnidirectional Imaging System: FIPPO と名付けた．FIPPO の性能は表 1 の最右列に示すとおりである．

図 4 に光学ユニットが形成する射影を示す．図はカメラ像面上の正方格子をパノラマシーン上へ投影した様子である．有効視野は方位方向におよそ $\pm 51.5^\circ$ ，縦方向におよそ $-31.8^\circ \sim +26.6^\circ$ である．また光学ユニットは装着者 (=撮影者) の視野を損なわない範囲で視点を下げて配置されている．その結果視点間の隔たりはおよそ 5.0cm となった．目標値である 4.4cm をわずかに満たしていないが，重要なのは隔たりによって生じる視線の不一致度である．後述する実験により， 5.0cm の隔たりでもほぼ目標を達成できることが確認されている．

4.2 機能的なスペック

FIPPO に用いたカメラは SANYO 製 Xacti DMX-HD2 である．このカメラは単体で $1280 \times 720\text{pixel}$ の

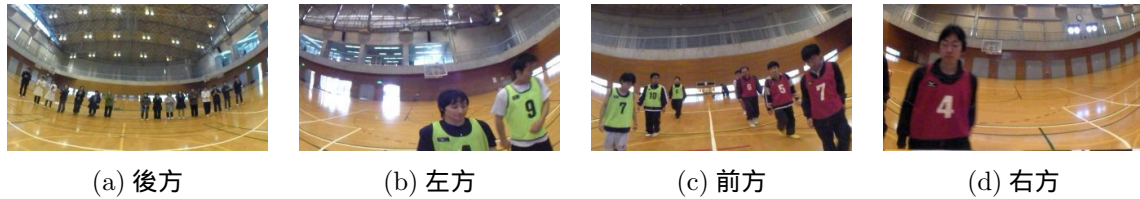


図 5: 撮影された映像例.



図 6: 映像歪みの補正を施して作成したパノラマ画像. 図 5 の (a)-(d) を元に処理した結果.

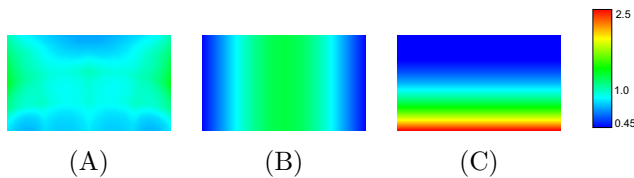


図 7: 正規化解能分布図. 図示範囲は方位方向 (横軸) に $\pm 45.0^\circ$, 仰俯角方向 (縦軸) に $-31.8^\circ \sim +26.6^\circ$ である. (A) FIPPO (var. = 0.012) (B) Multiple camera (var. = 0.077) (C) HyperOmni-Vision (var. = 0.42)

ハイビジョン映像をビデオレート (30fps) で撮影, 記録することができる. 追加の記録装置や電源線を必要としないため独立性の高い撮影システムを構成できた. バッテリー駆動による FIPPO の連続撮影時間はおよそ 85 分である. FIPPO の大きさは幅 40cm, 奥行き 40cm, 総重量は約 2.4kg であった.

4.3 試作機を用いたテスト撮影

FIPPO を頭部に装着し, 繁華街, 交差点, 動物園, 公園散策などの屋外シーンや屋内におけるスポーツシーンを実際に撮影した. 図 5 に屋内で行ったバスケットボールゲームシーンの収録例を示す. FIPPO は広視野に渡って均一な分解能を有しているが, 図 5 にも見られるように多少の幾何歪みが入力画像に生じる. 前後左右の入力映像を 1 本のパノラマ映像に結合するには, 幾何歪みの補正, さらにカメラ間の色補正を行った上で, 隣接する部分を繋げる必要がある. 図 6 に入力画像を繋いで生成したパノラマ映像例を示す. 本論文は撮影システムの提案を主とするため, つなぎあわせに関する手法の記述は省略する.

5 映像評価

5.1 全体解像度

全体解像度 (画素数) は有効視野にどれだけの画素が割り当てられているかで決定される. 隣接する光学ユニットによる重なり視野を評価対象から除き, 正しい解像度を求めた結果, 全体解像度はおよそ 284 万画素であった. FIPPO を構成する光学系では注目する方位や仰俯角によって割り当て画素数が変化する. 最小値と最大値を用いた範囲で表記すると, 方位方向に $3960 \sim 5024\text{pixel}$, 縦方向に $584 \sim 682\text{pixel}$ である.

5.2 分解能の均一さ

FIPPO の光学系における各パラメータからカメラ像面上の画素とシーン中の光線の対応は既知である. 分解能はシーン中における単位広がりには割り当てられている画素数で定義される. パノラマシーンを対象にした場合, 分解能 $R(\theta, h)$ は

$$R(\theta, h) = \frac{\delta u(\theta, h)\delta v(\theta, h)}{\delta\theta\delta h} \quad (6)$$

のように表すことができる, ここで, $h = \tan\phi$ である. また $u(\theta, h), v(\theta, h)$ は光線 (θ, h) に対応する画素を示す. 評価対象は分解能の絶対量ではなく分布であるので, 全体の平均値で除算した正規化解能を用いて分解能の均一さを評価する.

正規化解能の分布図を分散値と併せて図 7 に示す. FIPPO を構成する 4 台の光学ユニットは全て同じ構造であるので, 1 台による分解能分布を図示している. 従来手法に比べ FIPPO は有効視野全体にわたって均一な分解能を形成していることがわかる. 一方で全方位マルチカメラシステムでは仰俯角方向へは均一であるが, 方

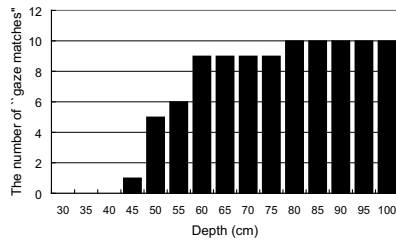


図 8: 視線の一致感が得られた距離のヒストグラム. 横軸は呈示顔画像を撮影した距離 D , 縦軸は視線一致感を得た被験者数を示す.

位の増加とともに分解能が減少する傾向にある. 双曲面ミラーを用いた全方位カメラはその逆で, 仰俯角方向への増減が激しいことが見てとれる. 各撮影システムにおける分散値も, FIPPO が従来手法よりも均一な分解能分布を形成していることを裏付けている.

5.3 視線一致度の評価

FIPPO によって撮影された映像が実際にどれだけの視線一致を与えられるか実験を通して評価した. 近距離状態での顔画像は 2 視点間の隔たり (2.1 節参照) の影響を大きく受けているので, 視聴者に視線の不一致感を与えてしまうが, 撮影距離が長くなればその感覚は徐々に減少する. まず, 撮影者と向かい合う共演者 (図 1) の距離 D を徐々に大きくしながら撮影を行い, 複数距離における共演者の顔画像列を得る. 続いて 10 人の被験者に対して, 撮影した顔画像列を撮影距離順に呈示し, どの時点で視線の不一致感が消失したかを調査した. 図 8 は実験結果を元に視線の一致感を得られた距離をヒストグラムで示したものである. 50cm 奥行き顔画像で半数の被験者が, 60cm で 9 割の被験者が視線の一致を感じている. 頭部搭載カメラによる視線一致の限界距離 1.4m (2.1 節参照) よりも大きく改善されており, 目標値である 50cm をほぼ達成している.

6 まとめ

我々は撮影者視点から見た全方位映像を撮影するための新たな装着型システムを開発した. 提案システムは, 記録装置を内蔵した小型カメラと特別に設計された曲面ミラーを組み合わせた独立性の高い構成で, 均一な分解能の全方位映像を撮影することができる. ミラーにより光学系の仮想視点を撮影者の視点に近づけているため, 近接した対象物に対しても高い視線一致感を得られる. さらに試作機を用いて屋外シーンを実際に撮影し, その実用性を確認した. 一方で様々なシーンに対応するには,

さらに小型軽量化が必要と考えている.

参考文献

- [1] K. Yamazawa, and H. Takemura, and N. Yokoya, “Telepresence system with an omnidirectional HD camera”, Proc. of ACCV2002, Vol. II, pp. 533–538, 2002.
- [2] Y. Onoe, K. Yamazawa, and N. Yokoya, “Telepresence by real-time view-dependent image generation from omnidirectional video streams”, Computer Vision and Image Understanding, Vol. 71, No. 2, pp. 154–165, 1998.
- [3] S. Ikeda, T. Sato, M. Kanbara, and N. Yokoya, “Immersive telepresence system with a locomotion interface using high-resolution omnidirectional videos”, Proc. IAPR Conf. on Machine Vision Applications, pp. 602–605, May 2005.
- [4] C. Cruz-Neria, D. J. Sandin, and T. A. DeFanti, “Surround-Screen Projectorion-Based Virtual Reality : The Design and Implementation of the CAVE”, Proc. of SIGGRAPH1993, pp. 135–142, 1993.
- [5] K. Yamazawa, Y. Yagi, and M. Yachida, “Omnidirectional Imaging with Hyperboloidal Projection”, Proc. of IEEE/RSJ Int. Conf. on Intelligent Robotics and Systems, Vol. 2, pp. 1029–1034, 1993.
- [6] Point Grey Research Inc.: “Ladybug2”, <http://www.ptgrey.com/products/ladybug2/>
- [7] 野呂影勇, “図説エルゴノミクス”, 日本規格協会, 1990.
- [8] S. M. Anstis, J. W. Mayhew, and T. Morley, “The perception of where a face or television portrait is looking”, American Journal of Psychology, 82(4), pp. 474–489, 1969.
- [9] R. Swaminathan, S. K. Nayer, and M. D. Grossberg, “Designing of Mirrors for catadioptric systems that minimize image error”, Proc. of OMNIVIS2004, 2004.
- [10] R. Swaminathan, “Focusing in Catadioptric Imaging System”, Proc. of OMNIVIS2004, pp. 1–7, 2007.