

複数のハンディカメラを利用した Diminished Reality

榎本 暁人[†] 斎藤 英雄[†]

[†] 慶應義塾大学 大学院理工学研究科 〒 223-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: [†]akihito@ozawa.ics.keio.ac.jp

あらまし 本論文では、複数のハンディカメラを用いた Diminished Reality システムを提案する。本システムでは、複数のハンディカメラが同じシーンを映しており、かつ見たい対象が障害物によって隠れてしまっている場合を想定し、異なる視点のカメラ画像を支援情報としてそれぞれのカメラの画像中から障害物を除去する。本システムでは、オンラインで各カメラの射影行列を求めるために、ARTag マーカを利用する。一方、対象シーンを平面近似し、近似された平面領域に対する各カメラ間の Homography をオンラインで得られる射影行列により算出する。そして、算出された Homography を用いて平面射影変換された画像をユーザ視点画像と合成することで、平面として近似されない障害物領域が除去される。本システムの有効性を確認するために、全ての自由に動いているカメラにおいて、不要な障害物を除去した映像をオンラインで合成可能なことを示した。

キーワード 拡張現実感, Diminished Reality, Homography, ARTag, 複数カメラ

Diminished Reality using Multiple Handheld Camera

Akihito ENOMOTO[†] and Hideo SAITO[†]

[†] Graduate School of Science and Technology, Keio University 3-14-1, Hiyoshi, Kouhoku-ku, Yokohama, Kanagawa, 223-8522 Japan

E-mail: [†]akihito@ozawa.ics.keio.ac.jp

Abstract In this paper, we present a system for Diminished Reality with multiple handheld camera system. In this paper, we assume the situation that the same scene is captured with multiple handheld cameras, but some objects occlude the scene. In such case, we propose a method for diminishing the occluding objects from each camera by the support of the different cameras that capture the same scene at the different viewpoints. In the proposed method, we use the AR-Tag marker to calibrate the multiple cameras, so online process can be possible. By the use of AR tag, we compute a homography for between each camera's imaging plane and the objective scene that is approximated as a planar in this research. Based on the homography, we warp the planar area for synthesizing the viewer's image that includes only the objective scene without the occluding object that cannot be approximated as a planar area. We will demonstrate that the viewer's image without the occluding object can be synthesized for every camera on-line, even though all the cameras are freely moving for all cameras each camera that relates.

Key words Augmented Reality, Diminished Reality, Homography, ARTag, Multiple Camera

1. はじめに

近年、コンピュータビジョンの分野において、さまざまなシーンで複数カメラを用いた研究が行われている。複数のカメラを使うことで、1つのカメラに比べてより多くの情報を手に入れることが出来る。Augmented Reality(AR)/Mixed Reality(MR)の研究分野においても、複数カメラを用いたさまざまな研究が成されてきた。佐藤ら [1] は、頭上に設置された固定カメラを用いて、ユーザの頭部の位置・姿勢を追跡することでユーザカメラの外部パラメータを修正する手法を提案した。Julien ら [2]

は、複数の固定カメラを用いて、AR アプリケーションのための幾何学的キャリブレーション・光学的キャリブレーションを行うシステムを提案した。Julien らのシステムにおいて、ユーザカメラ以外のカメラは、ユーザカメラの幾何学的キャリブレーション・光学的キャリブレーションを行うために用いられる。George ら [3] は固定カメラを用いて、タブレット PC の取り付けられた LED の追跡を行った。LED を追跡することで、タブレット PC に設置されたカメラの位置・姿勢を正確に推定できる。市原ら [4] はロードサイドカメラを用いて NaviView と呼ばれる運転支援システムを提案した。上記に述べたように、

複数カメラは様々な研究で用いられてきたが、複数カメラを用いたシステムの大部分が以下の制約を有している。1) 固定カメラが必ず1台以上存在している。2) 固定カメラはユーザカメラの支援に用いられる。そこで本論文では、我々は全てのカメラが固定されておらず、かつ全てのカメラがユーザカメラになり得るという複数カメラシステムを提案する。固定カメラが必要であるという制約が必要ないので、我々の提案システムは従来の複数カメラシステムよりも柔軟かつ広範囲の様々なアプリケーションに用いることができる。本論文において、我々は固定カメラが必要でない複数カメラシステムを用いた Diminished Reality システムを提案する。

Augmented Reality(AR) が仮想物体をあたかも実環境に存在するかのように重ね合わせて表示する技術であるのに対して、Diminished Reality(DR) とは適切な背景画像を障害物領域に重ね合わせるにより画像中から障害物を取り除く技術である。Diminished Reality の従来研究は大きく分けて2つのアプローチに分けることが出来る。

1つ目のアプローチは、1台のカメラで撮影された動画の時系列情報を用いる手法である。Lepetit ら [7] は、ユーザが指定した境界線を追跡してオクルージョンを検出することでシーン中の障害物を取り除く手法を提案した。Mann, Fun [5] は、入力シーンから動画像列に撮影される平面物体を取り除き、別の平面テクスチャに置き換える手法を提案した。また、Wang, Adelson [6] は動画像列を複数のレイヤーに分割し、一つのレイヤーを取り除く場合には同じ動画をレンダリングする手法を提案した。Diminished Reality という言葉は用いていないが、横井ら [8] は講義ビデオの生成において、映像中から講師を除去する手法を提案した。横井らの手法では、まず現フレームでの講師領域を検出し、その検出された講師領域に1フレーム前の画像を用いることで講師の除去を実現した。2つ目のアプローチは、複数の位置に設置したカメラに撮影された画像を用いる手法である。Zokai ら [9] は、大きく離れた3視点から撮影された静止画像の投影モデルを用いて適切な背景を生成することでユーザが指定した障害物の除去を行う手法を提案した。3つのカメラの役割は中央カメラがユーザカメラ、他の2台のカメラが背景画像生成用のカメラである。Zokai らの手法は背景領域の3次元復元を行うため、背景が非平面の場合でも扱うことが出来た。実際に、Zokai らの手法は工場の再設計において用いられ、有用性を証明した。

本研究では、複数のハンディカメラが同じシーンを映しており、かつ見たい対象が障害物によって隠れてしまっている状況を想定し、異なる視点のカメラ画像を支援情報として各カメラの画像中から障害物を除去するシステムを提案する。本システムは Zokai らのアプローチと同様に、異なる視点から撮影された複数の画像を使用する。Zokai らの手法において、1台のカメラは対象シーン撮影用のユーザカメラ、他の2台は背景画像撮影カメラというようにそれぞれのカメラの役割は分かっていた。我々の提案手法においては、全てのカメラがユーザカメラになり、同時に背景画像撮影カメラにもなる。複数の可動カメラを用いてオンラインでの Diminished Reality を実現するため

に、我々は ARTag [10] を用いてオンラインキャリブレーションを行う。オンライン処理を行える利点として、ユーザの自由な視点移動、障害物の移動、対象シーンの変化に対応して処理を行えるということが上げられる。

最後に、本稿の構成について述べる。まず2章では提案システムの概要について記述する。3章では、オンラインでの Diminished Reality を実現するための提案手法について述べる。4章では、本システムを用いて障害物を除去した結果を示すことで本システムの有効性を示す。そして5章で結論を述べる。

2. 提案システム

現在、あらゆる環境中に多くのカメラが存在している。例えば、ほとんどの人がカメラ付携帯電話を持っている。本論文では、図1に示されるように、複数の人が持つカメラを利用することで Diminished Reality システムを構築する。全てのカメラの画像を共有することで、全てのユーザ視点において障害物が除去された画像を生成することが出来る。このシステムを応用することで、カメラ付携帯電話を持っている全ての人が障害物のない対象シーンを見ることが出来る。例えば、美術館において有名な絵画が前にいる人々によって隠れてしまっても、本システムを用いれば障害物のない絵画の映像を手に入れることができるのである。

このコンセプトに基づいて、我々は図2に示すようなプロトタイプの実験システムを構築した。このシステムは複数の可動カメラから成り立つ。全てのカメラは異なるユーザが持ち、異なる視点から同一のシーンを撮影する。実験システムにおいて、我々は異なる PC に接続された複数の web カメラを用いる。それぞれの PC はネットワークに繋がっているため、ネットワークを通じて撮影された画像データの共有が可能である。ユーザは環境中に配置された ARTag マーカが検出される範囲で自由にカメラを動かすことで、オンラインでのキャリブレーションが可能となる。ARTag マーカを対象平面と同一平面上に配置した場合は、オフラインでの手動作業が必要でない。例えば対象シーンが平面ではなくとも、カメラから遠景のシーンを対象平面と近似することで処理を行うことが可能である。対象シーンと ARTag マーカが配置された平面が同一平面上でない場合、前処理としてオフラインで対象シーンの3次元位置を取得することで Diminished Reality を行うことが可能である。我々の提案システムでは3台以上のカメラの使用が可能であるが、今回は図2に示されるように、最小のシステム構成として3台のカメラを用いて提案システムの有用性を証明する。

このシステムにおいて、我々は全てのカメラにおいて他の2台のカメラの画像を支援情報として障害物の除去を行う。本論文で、我々は障害物除去が行われるカメラをユーザカメラ、一方、ユーザカメラの支援に用いられる他の2台のカメラをサポートカメラと呼ぶ。

3. 障害物の除去

ここでは、提案システムを実現するために本研究で提案する手法について述べる。図3で示されるように、我々の提案手法

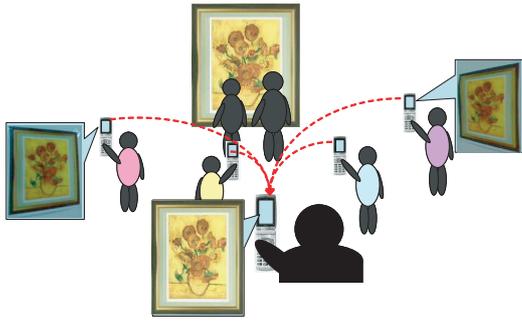


図1 提案システム

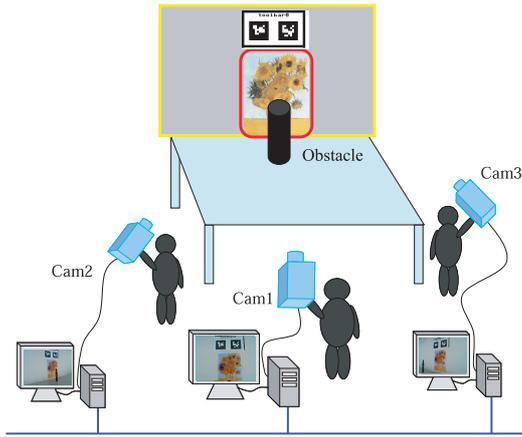


図2 プロトタイプの実験システム

は大きく分けて以下の2つの処理から成り立つ。

- (1) 射影変換画像の生成
- (2) ブレンド処理

1つ目の処理である射影変換画像の生成では、対象シーンは平面であると近似することで撮影された画像を他のカメラ視点に変換する。平面射影変換を行うために、ARTag マーカを用いてカメラのオンラインキャリブレーションを行う。

2つ目の処理であるブレンド処理では、撮影された画像と他カメラの射影変換画像を統合する。我々はブレンド処理において、対象平面以外の領域すなわち障害物部分を除去する。以下、それぞれの処理について詳しく述べる。

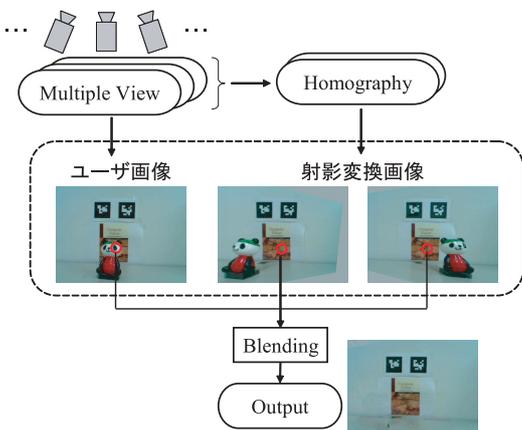


図3 処理の流れ

3.1 射影変換画像の生成

3.1.1 ARTag マーカによる射影行列算出

我々はカメラの射影行列 P を算出するために ARTag と呼ばれるソフトウェアを利用する。ARTag とは画像処理によって、カメラの入力画像中からマーカを検出することで、オンラインでカメラの射影行列 P を高速に精度良く算出することのできるソフトウェアである。ARTag により算出される射影行列は図4に示されるような世界座標系を持つ。



(a) 対象平面中に配置 (b) 垂直な平面に配置
図4 世界座標系

3.1.2 対象平面に対する Homography の算出

マーカ検出により得られた射影行列 P を用いることで対象平面に対する Homography を算出する。対象平面と画像平面間の Homography が算出することができれば、対象平面に対するカメラ間での Homography は式(1)のようにして求まる。図5で示される対象平面上でのある点 M と画像 A, 画像 B 上での対応点 m_A, m_B 間での Homography がそれぞれ H_A, H_B で表されたとする。この時、画像 A から画像 B への Homography(H) は一般に以下の式で表される。

$$H = H_B^{-1} H_A \quad (1)$$

次にカメラ間の Homography を求めるために必要な対象平面と画像平面間の Homography の算出法について述べる。図4(a)で示されるように、マーカを対象平面上に配置した場合は世界座標系の $Z=0$ 平面が対象平面となる。射影行列 P より $Z=0$ の時を考えると

$$\begin{aligned} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} &\simeq P \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} \simeq \begin{pmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \\ &\simeq H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \end{aligned} \quad (2)$$

となり、 3×4 の射影行列 P を 3×3 の行列 H で表すことが出来る。すなわち、この式は2次元座標から2次元座標への変換となる。よって、世界座標系の $Z=0$ 平面から画像座標系への Homography が算出できる。

3.2 ブレンド処理

まず、図6は3枚の画像を全て等しい割合(1/3ずつ)でブレンドした結果である。3枚の画像中の障害物がそれぞれうっすら残ってしまっている。図6で示される様に、画素毎に適切な割合でそれぞれの明度値をブレンドしなければ、障害物が画像

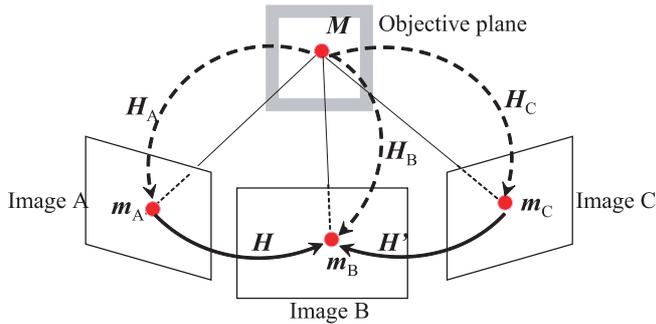


図5 カメラ間の Homography

中に残ってしまう．ここでは，以下で障害物の除去をするために行われるブレンド処理について詳述する．ブレンド処理とは，ユーザ視点画像と2枚の射影変換画像間で対応する画素の明度値の差に応じて，明度値を混ぜる処理である．もし対応する3つの画素の明度値がほぼ等しい場合，この場合はユーザ視点画像においては対象平面領域であると考えられるのでユーザ視点画像の明度値を選択する．次に，明度値が等しくない場合，この場合は明度値が最も近い2つの画素を選択し，大きく異なる明度値は無視する．ユーザ視点画像の明度値のみが他の2つの明度値と大きく異なる場合，この場合はユーザ視点画像においては障害物領域であると考えられるので，他の2つの明度値を選択する．

便宜上，画像座標 (i, j) の位置でのユーザ視点画像の明度値を p_V ，射影変換画像 n の明度値を p_n と定義する．ここで，画像座標 (i, j) の位置での出力画像の明度値を q とすると，

$$q = p_V \quad (p_V \approx p_1 \approx p_2) \quad (3)$$

$$q = (p_1 + p_2) / 2 \quad (p_V \neq p_1 \approx p_2) \quad (4)$$

となる．ユーザ視点画像の障害物領域では，ユーザ視点画像の明度値は他の2枚の射影変換画像の明度値とは大きく異なる．よって，これらの式を用いることでユーザ視点画像の障害物領域では射影変換画像の明度値を用いることとなるので，障害物は出力画像において除去されるのである．



図6 ブレンド処理の失敗例

3.3 対象平面以外への ARTag マーカの配置

前項で述べたように，ARTag マーカを対象平面上に配置した場合は，対象平面に対するカメラ間の Homography がオンラインで算出できる．しかしながら，たとえ ARTag マーカを対象平面上に配置することが出来なくても，オフライン処理によって ARTag マーカと対象平面の位置関係を取得することで障害物除去の処理を行うことが出来る．ARTag マーカを対象平面以



図7 Blending process

外の位置に設置した場合は，オフライン段階において対象平面の3次元位置を指定する必要がある．

ここでは，マーカを対象平面に対して垂直な平面に配置した場合を考える．マーカを垂直な平面に配置した場合は，世界座標系の $X=d$ の平面が対象平面となる．この場合，対象平面に対する Homography は以下の式によりもとまる．

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \approx P \begin{pmatrix} d \\ Y \\ Z \\ 1 \end{pmatrix} \approx \begin{pmatrix} p_{12} & p_{13} & p_{14} + dp_{11} \\ p_{22} & p_{23} & p_{24} + dp_{21} \\ p_{32} & p_{33} & p_{34} + dp_{31} \end{pmatrix} \begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix} \quad (5)$$

$$\approx H \begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix}$$

ここで，射影行列は2次元座標から2次元座標への変換となり，すなわちこの変換は $X=d$ と画像平面間での Homography となる． d はオフライン処理において3枚の画像間での対応点を手動で指定することによりもとまる．

4. 実験・結果

ここでは，実験を通して提案システムの有用性を証明する．マーカを対象平面と同一平面上に配置した場合，マーカを垂直な平面に配置した場合の2つのマーカ配置について，提案手法を用いて障害物の除去が実現出来ているかを調べた．本実験は最小のシステム構成としてカメラ3台を用いて行った．本実験で用いた PC のスペックは以下の通りである．

- CPU : Intel Pentium4 Processor 3.0 GHz
- メモリ: 1024 MB
- OS : Windows XP

4.1 マーカを対象平面上に配置

マーカを図8のように対象平面上に配置した場合について，対象平面の前に置かれた障害物を本手法を用いて除去した．



図8 マーカの配置位置

まず，実験1として3台のカメラの内，中央のカメラをユーザカメラとして処理を行った時の結果を図9に示す．図9(a)においては障害物によって対象平面が隠れてしまっているが，図9(b)の出力では障害物が除去され，見たい対象平面の情報が得

られる事が分かる .



(a) 入力 (b) 出力

図9 実験1の結果

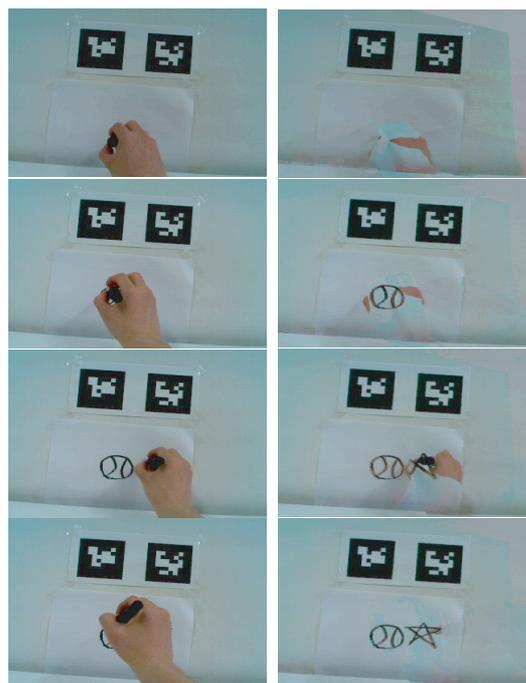
また、実験1において3台のカメラ全てをユーザカメラとした場合の結果を図10に示す。図10は上から順に、中央・左・右カメラをユーザカメラとした時の入力と出力であるが、どの位置からでも障害物が除去されている。よって、ユーザは自分の見たいどの位置からでも障害物が除去された画像を取得出来る事が分かる。



(a) 入力 (b) 出力

図10 実験1-2

実験2では対象平面上の白紙に人が絵を書いているという状況である。この時ユーザカメラにおいては人の手が障害物となってしまう書いた絵(野球ボールと星)が隠れてしまっている。しかし、提案手法を用いることで図9(b)で示される様に人の手を消し、書いた絵を見ることが出来た。出力画像には一部手が消えずに残ってしまっている部分があるが、これはユーザ視点画像中の障害物領域と投影画像中の障害物領域が重なってしまったためブレンド処理がうまくいかずに障害物が残ってしまったのと考えられる。



(a) 入力 (b) 出力

図11 実験2の結果

次に実験3として、今までは見たい対象が完全な平面のみであったが、対象が平面ではなく図12のような立体の場合に対して本手法を用いて処理を行った。結果を図13に示す。



図12 遮蔽状態

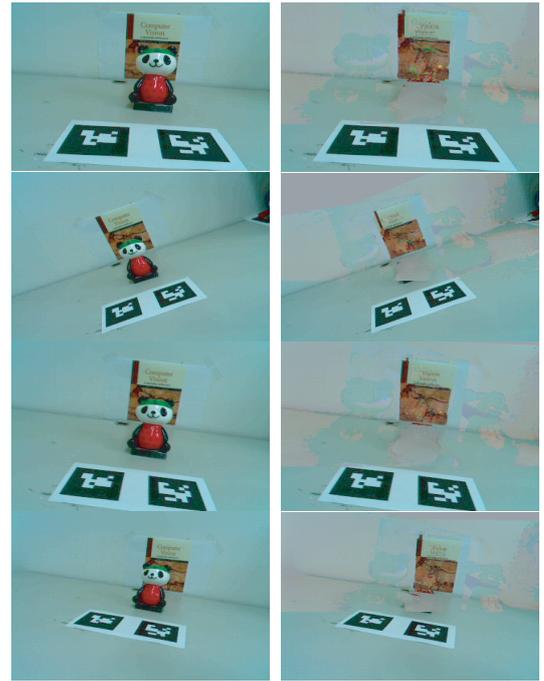
図13より、対象が平面の時と比べて出力画像における見たい対象物体はボケてしまっている。このようにボケてしまった理由として、マーカを対象平面と垂直な平面に配置した場合は世界座標系 $X=d$ の位置に平面が存在していると仮定して処理を行っている。しかしながら、実際は立体なため平面同士の対応付けである Homography を用いて処理を行った際に誤差が生じてしまうためである。



(a) 入力

(b) 出力

図 13 実験 3 の結果



(a) 入力

(b) 出力

図 14 実験 4 の結果

4.2 マーカを対象平面と垂直な平面上に配置

マーカを対象平面と垂直な平面上に配置した場合について本手法の処理を行った。実験 4 では、実験 1・実験 2 と同様に対象が平面の場合についての結果である。マーカを対象平面と垂直な平面上に配置した場合は、オフライン処理を行い対象平面の 3 次元位置を算出する必要がある。今回オフライン処理でユーザが 2 点、対応点を指定してオフライン処理を行うことで対象平面の 3 次元位置を取得した。実験 4 の結果を図 14 に示す。

実験 4 の結果より、マーカを対象平面と垂直な平面上に配置した場合でも入力画像における障害物が除去されていることが分かる。

5. おわりに

我々は今回、複数のハンディカメラを用いた Diminished Reality システムを提案した。我々の提案システムにおいて、全てのカメラは固定である必要がなく、また全てのカメラにおいて障害物の除去が行われる。我々は 2 つのマーカ配置の場合における障害物の除去を実験を通して行った。ARTag マーカを使うことで、高速なキャリブレーションが可能となるので、我々の提案手法はオンラインでの動作が可能となる。実験結果は、提案手法が障害物が動いた場合、カメラが動いた場合、対象平面が時間的に変化した場合でも障害物除去が行われていることを示している。今回の処理では領域を平面近似して処理を行っているため、対象が立体の時は対象にボケが生じてしまったが、今後カメラの台数を増やすなどして立体にも対応できるように処理を改良する予定である。

文 献

- [1] Kiyohide Satoh, Shinji Uchiyama, and Hiroyuki Yamamoto, A Head Tracking Method Using Bird 's-Eye View Camera and Gyroscope, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality(2004)
- [2] Julien Pilet, Andreas Geiger, Pascal Laguerre, Vincent Lepetit, Pascal Fua, An All-In-One Solution to Geometric and Photometric Calibration, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality(2006)
- [3] Georg Klein, Tom Drummond, Sensor Fusion and Occlusion Refinement for Tablet-based AR, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality(2004)
- [4] Eitaro ICHIHARA, Yuichi OHTA, NaviView: Visual Assistance Using Roadside Cameras- Evaluation of Virtual Views -, IEEE Intelligent Transportation Systems Conference Proceedings 2000 Dearborn
- [5] S. Mann and J. Fung, Videorbits on eye tap devices for deliberately diminished reality or altering the usual perception of rigid planar patches of a real world scene, International Symposium on Mixed Reality (ISMR2001)
- [6] Adelson and Wang, Representing moving images with layers, IEEE Transactions on Image Processing Special Issue: Image Sequence Compression, September 1994.
- [7] V. Lepetit and M.-O. Berger, A semi-automatic method for resolving occlusion in augmented reality, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 00), June 2000.
- [8] Takao Yokoi, Hironobu Fujiyoshi, Generating a Time Shrunk Lecture Video by Event Detection, ICME2006 : IEEE International Conference on MULTIMEDIA and EXPO
- [9] Siavash Zokai, Julien Esteve, Yakup Genc and Nassir Navab, Multi-view Paraperspective Projection Model for Diminished Reality, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2003).
- [10] Mark Fiala, The SQUASH 1000 Tangible User Interface System, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2005)