

## 第 6 章

### 新規映像制作技術への応用

#### 6.1 はじめに

近年、ハイビジョン放送やデジタル放送の実用化、インターネットによる映像配信の普及により、映像メディアの高品質化や多様化、多チャンネル化が急速に進んでいる。今後、臨場感のある映像コンテンツへのニーズはますます増加し、魅力的な映像を効率的に制作する技術革新が必要となる。

本章では、Axi-Vision カメラにより次世代の映像制作技術を開拓することを目指し、距離情報を利用した映像合成技術や 3 次元モデリング、立体映像システムなどへの応用の可能性を検証する。まず、第 1 の応用として第 6.2 節では、距離情報を用いた映像合成技術の確立を目指し、Axi-Vision カメラによる映像合成の基本性能を調べ、奥行き距離を考慮した 3 次元的な映像合成手法を開発するとともに、バーチャルスタジオにおける番組制作に実用化する。次に、第 2 の応用として第 6.3 節では、画像蓄積効果によるノイズ低減手法を提案し、距離検出分解能の向上による被写体の 3 次元モデリングとその映像制作への応用の可能性を調べる。最後に、第 3 の応用として第 6.4 節では、立体映像システムへの応用を目指し、距離画像をもとに奥行き標準化画像を生成するとともに、立体表示装置を試作する。Axi-Vision カメラと立体表示装置により、実時間の立体撮像・表示実験を行い、将来の立体映像システムとしての可能性を検証する。

#### 6.2 距離情報を用いた映像合成手法

本節では、Axi-Vision カメラで得られた距離情報を用いた映像抽出と合成技術の確立を目指す。現在、一般に映像合成にはクロマキー法[1]が用いられているが、カラー映像信号の色成分における色相の差を利用し映像合成するため、ブルーの背景(ブルー

バック)が必要であり(図 6.1(a))、その背景を高輝度かつ均一な輝度で照明する煩雑な作業も伴う。また、出演者の服装の色が制限され、被写体へのブルーバックからの反射光の影響や切出しエッジ部分への色漏れなどの課題もある。

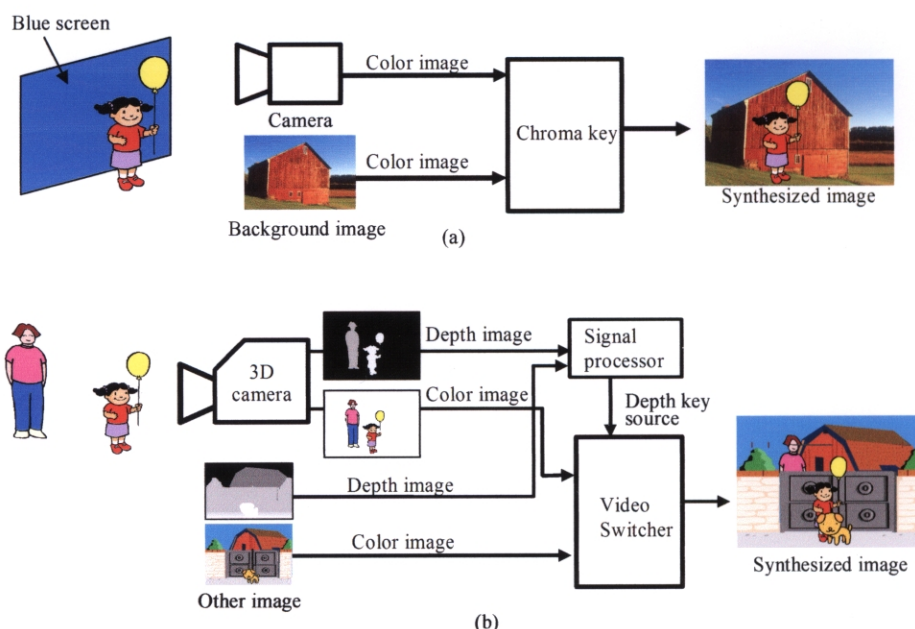


図 6.1 映像合成手法 (a) クロマキー手法 (b) 距離情報を用いた映像合成手法

一方、距離情報を用いた映像合成法(図 6.1(b))では、被写体のカラー映像とともに取得された距離情報をもとに映像抽出キー信号を生成し、他の映像と合成する。この合成法の特長は以下の様になる。

- ① ブルーバックが不要となり、クロマキー法の課題を解決できる。
- ② 任意の距離に位置する被写体の映像のみを抽出し、他の映像と合成できる。
- ③ 奥行き距離を考慮した3次元的な映像合成が可能。

被写体の距離情報を活用することで、従来技術の課題を解決するとともに、臨場感や一体感のある映像表現が可能となる。さらに、映像制作設備の規模や作業量、コストの軽減効果も期待できる。

従来、距離情報を用いた映像合成手法の提案はあったが[2-5]、ハイビジョンクラスの高品質な映像制作に実用化できるだけの精細度とリアルタイム性を兼ね備えた距離検出技術はなかった。これに対し、本研究によるハイビジョン Axi-Vision カメラの開発により、距離情報を用いた映像合成技術を実際の高品質なコンテンツ制作に応用できる可能性が得られた[6]。

次節では、距離情報をもとにした映像合成法(Depth-key)の基本特性を調べるとともに、合成する被写体間の前後関係を考慮した3次元的映像合成手法を確立する。さらに、これらの技術をバーチャルスタジオに導入し、実際の番組制作への応用を図る。

## 6.2.1 Depth-key

Axi-Vision カメラにより得られた距離情報をもとにした Depth-key 法の基本性能を実験で確かめる。図 6.2(a)に示すように、カメラに対し 45 度傾斜したパネルを被写体とし、強度変調光の変調数を 43 MHz、シャッター時間を 2 ns とし、測定レンジ中心をカメラから 2.4 m の距離に設定し撮影した。被写体の映像とともに得られた距離映像信号  $I(x,y)$  (図 6.2(b))に対し、信号レベルの上限閾値  $T_1$  と下限閾値  $T_2$  で 2 値化したキー信号を生成後、ビデオスイッチャーで被写体のカラー映像(図 6.2(c)) から特定の距離範囲の被写体映像のみ(図 6.2(d))を抽出した。

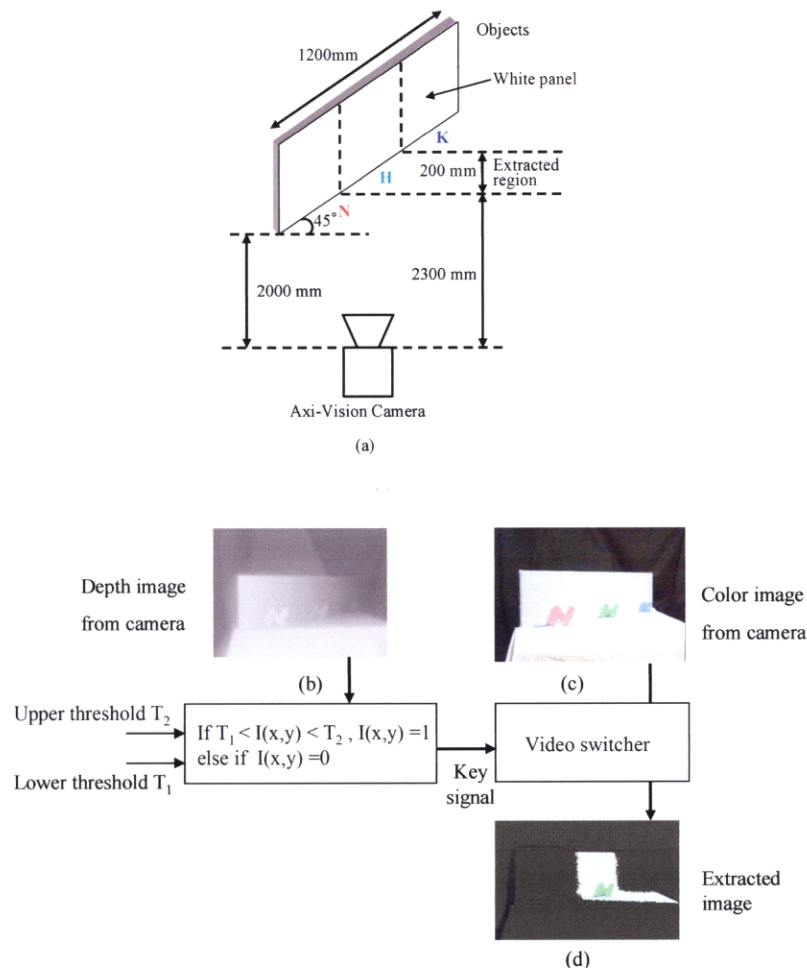
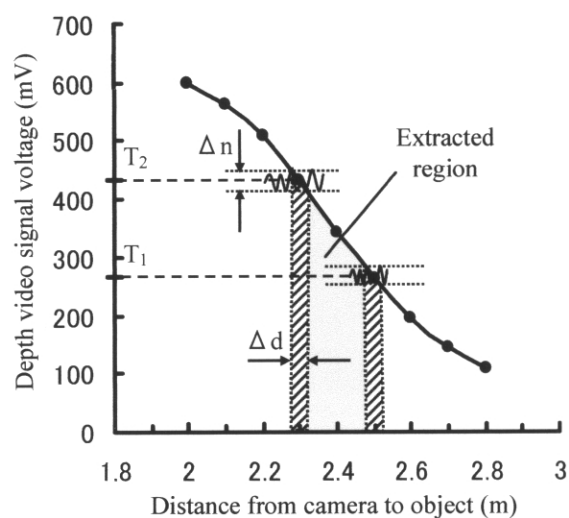
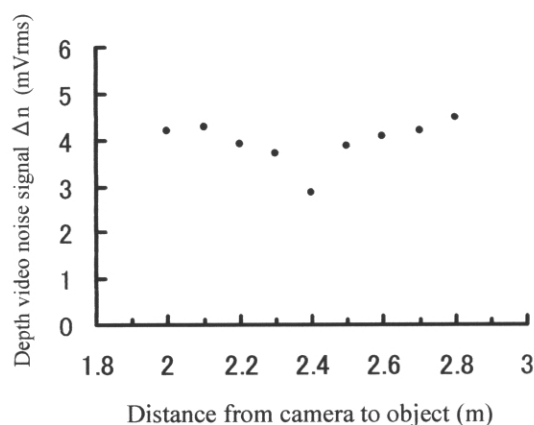


図 6.2 距離情報をもとにした映像抽出 (a) カメラと被写体の配置 (b) 距離画像 (c) カラー画像 (d) 抽出画像



(a)



(b)

図 6.3 出力距離信号 (a) 信号レベル (b) ノイズ信号

閾値  $T_1$ 、 $T_2$  の設定は、図 6.3(a) に示すようにカメラから被写体までの距離と出力距離信号レベルの関係より決定される。この実験では、カメラからの距離が 2.3 m から 2.5 m の範囲の被写体を抽出するために、 $T_1$  および  $T_2$  の閾値を、映像信号レベル 262 mV および 434 mV にそれぞれ設定した。

図 6.2(d) の抽出画像のエッジ部分にノイズがみられるが、これは図 6.3(a) に示すように、距離画像の検出ノイズ  $\Delta n$  に相当する距離検出の揺らぎ成分  $\Delta d$  が原因である。距離画像のノイズ成分  $\Delta n$  (mVrms) の測定結果を図 6.3(b) に示す。被写体のカメラからの距離が 2.3 m の場合、距離検出のノイズ信号は 3.7 mVrms、また距離 2.5 m の場合、3.9 mVrms である。この距離検出ノイズは、測定レンジ中心  $d=2.4$  m でもっとも少なく、測定レンジ周辺で増加している。この理由は、測定レンジ周辺部では、距離算出のも



ととなる増加、または減少変調光時の受光量が減少することで撮影映像の SN 比が低下し、第 4 章の(4.9)式からわかるように距離検出ノイズ成分が増加するためである。また、測定レンジの位置は、第 3 章の(3.7)式で示したように、シャッタータイミングで制御できるため、被写体の切出し閾値が測定レンジ中心付近となるように設定することで、抽出画像のエッジ部分のノイズの影響は最小に抑えることができる。

距離情報による被写体映像の抽出例を図 6.4 に示す。カメラで得られる距離画像(図 6.4(b))よりキー信号を生成し、各奥行き距離の被写体のみの画像(図 6.4(c)~(e))を抽出した。撮影時のカメラから人物までの距離は 2.5 m、カメラから背後のセットまでは 3.7 m である。切出しの閾値を変えることで、人物より遠い背景のみの映像(図 6.4(c))や、人物のみの映像(図 6.4(d))また、カメラに近い腕の部分だけの映像(図 6.4(e))を選択的に抽出できる。被写体がカメラに近すぎる場合、近赤外光の受光量が多く、CCD カメラでの撮影信号が飽和し、良好な距離画像が撮影できない場合もあるが、被写体距離に応じて I.I.の感度や照射光量を調整することで改善できる。

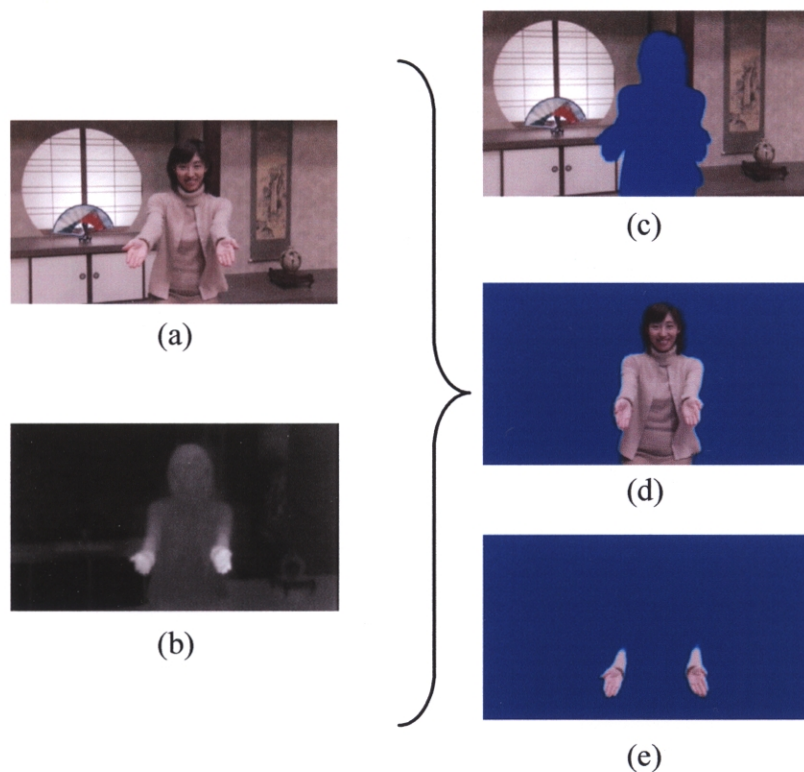


図 6.4 距離情報による映像抽出例 (a) 被写体の映像 (b) 距離画像 (c) 遠い被写体のみの抽出映像 (d) 人物のみの抽出映像 (e) 最もカメラに近い被写体のみの抽出映像

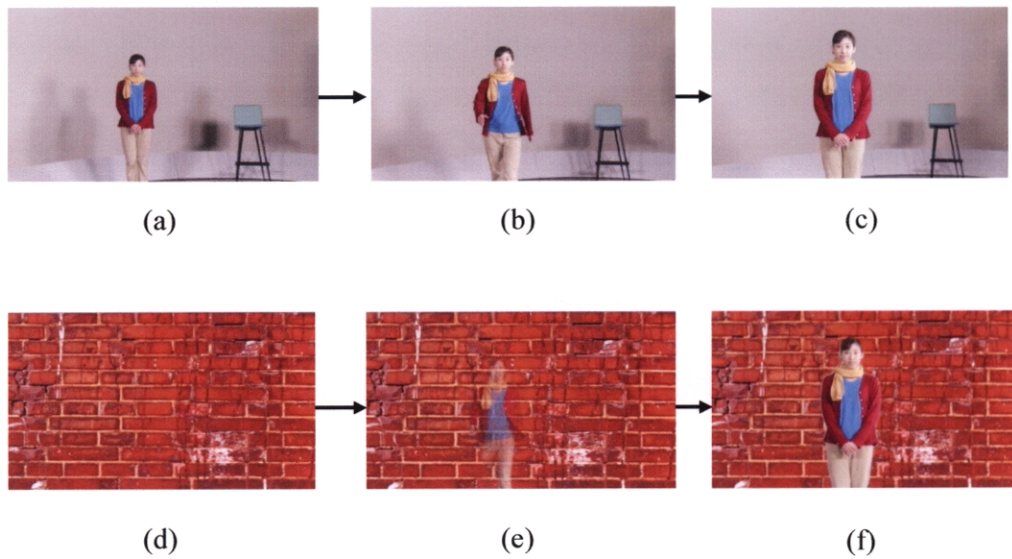


図 6.5 距離情報による映像合成実験 (a)～(c) 合成前の映像 (d)～(f) CG (レンガの壁の画像)合成後の映像

また、図 6.5 に Depth-key による実写動画像と CG との合成例を示す。奥行き距離約 7 m のスタジオ内で、カメラから 3 m の距離範囲内の被写体のみの映像を抽出し、CG のレンガの壁の画像と合成する設定とした。図 6.5(a)～(c)は合成前の映像、図 6.5(d)～(f)は合成後の映像を示している。人物が設定閾値より奥の場合は、CG 上に人物は合成されず(図 6.5(d))、カメラに人物が近づき、閾値を超えると人物が CG と合成され始める(図 6.5(e))。この閾値付近では距離検出ノイズが合成映像の画質に大きく影響するが、合成キー信号のエッジをソフトにしたり、合成映像のゲイン値を調整するなどの工夫で改善効果がみられる。被写体がさらにカメラに近づくと、完全に CG を背景とした合成映像が得られる(図 6.5(f))。このように、ブルーバックなしのスタジオでも、動く被写体に対し Depth-key による映像合成がリアルタイムで可能であることが実証できた。

## 6.2.2 3次元映像合成技術

実写映像と CG 映像の距離情報を実時間で比較合成することで、前後関係を考慮した 3 次元的な映像合成が実現できる。その合成の過程を図 6.6 に示す。被写体と CG キャラクタの相対的な前後関係は、被写体の距離画像  $D_1(x, y)$  (図 6.6(a))と CG の距離画像  $D_2(x, y)$  (図 6.6(c))の差分より得られる。差分  $\Delta D(x, y) = D_2(x, y) - D_1(x, y)$  は、被写体と CG の相対距離を与え、 $\Delta D(x, y) > 0$  の場合、CG は実写の被写体よりもカメ



ラに近く、反対に  $\Delta D(x, y) < 0$  の場合は、その画素は被写体が CG よりカメラに近いことを表している。図 6.6(e)はこの差分画像にセッアップ輝度信号を加えたものであり、画像の濃淡で相対的な前後関係を表した画像である。さらに、この図 6.6(e)の画像を 50%映像レベルで 2 値化したのが、図 6.6(f)に示す合成用のキー信号であり、実写映像よりカメラに近い CG 映像の抽出に使用される。すなわちキー信号の黒部分には、実写映像が表示され、白い部分には CG 映像が表示される。図 6.6(g)に示す合成映像では、人物より遠い距離の CG キャラクタは出演者の背後に隠れ、近い CG キャラクタは出演者の手前に表示されている。

クロマキー法による映像合成では、前後関係を考慮することなく、抽出画像は単に他の映像に上書き合成されるのみであるが、距離情報を使った合成では、前後関係を考慮に入れた 3 次元的な合成が可能となり、実写と CG があたかも同じ空間にいるかのような一体感のある合成映像が実時間で制作可能となった[7]。

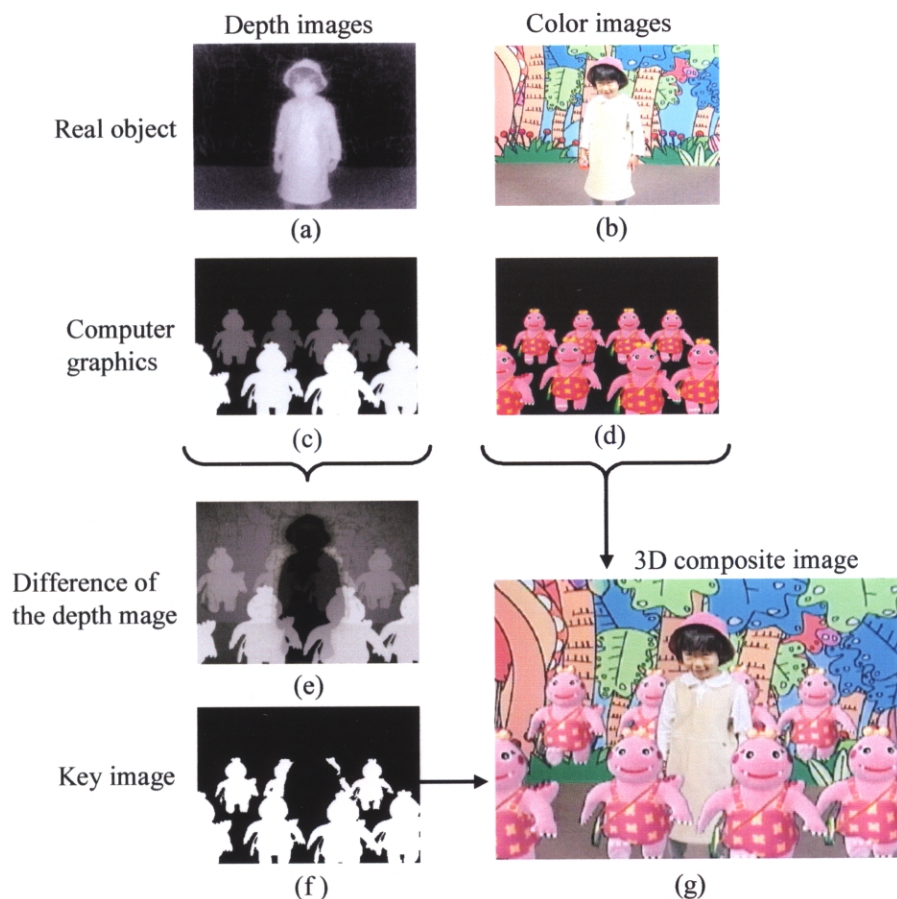


図 6.6 実写と CG の 3 次元映像合成 (a) 実写の距離映像 (b) 実写のカラー映像 (c) CG の距離映像 (d) CG のカラー映像 (e) 実写と CG の距離の差分映像 (f) 合成キー信号 (g) 3 次元合成映像

## 6.2.3 バーチャルスタジオへの導入と番組応用

### (1) バーチャルスタジオ

Axi-Vision カメラと Depth-key システムによる新たなバーチャルスタジオを開発した。バーチャルスタジオとは、CG 制作したスタジオセット映像を出演者の映像と合成するものである。現実にはありえないスタジオセットを CG 制作することで、多彩な映像演出を行うとともに、スタジオセットの製作コストの低減や、コンテンツ制作規模の縮小化などが可能である。近年、バーチャルスタジオでは、カメラの動きに応じて、背景 CG をリアルタイム生成し合成するシンセビジョン[1]や、事前に収録した実写映像を背景として合成する実写ベースバーチャルスタジオ[8,9]が開発されている。しかし、これらのバーチャルセットでも、映像合成には従来のクロマキー法が用いられていた。そこで、Axi-Vision カメラによる距離合成技術を導入し、ブルーバックが不要な新たな実空間ベースバーチャルスタジオを開発した[10]。

図 6.7 に Axi-Vision カメラを使ったバーチャルスタジオの概要を示す。まず、事前に製作された舞台セットにおいて、固定位置に置かれたカメラで全周囲の映像を撮影し、パソコンのハードディスクに収録する(図 6.7(a))。この映像撮影時には、映像と同期して、エンコーダ付のカメラ雲台より、カメラのパンとチルトデータ、カメラレンズのズームとフォーカスデータを同時に取得し記録しておき、超高精細な全天周画像(8192×8192 画素、2 枚)を作成する。

本番映像収録時には、出演者の演技に必要なセット部分のみを実セットとして作製し(図 6.7(b))、セットが無い部分は全天周画像よりカメラデータをもとにパソコンで画像を生成して合成する。合成の際は、Axi-Vision カメラで得られた距離情報を元に Depth-key 法で合成するため、実セット以外の部分をすべてブルースクリーンで覆う必要がなく、360 度あらゆる方向の撮影も可能となり、実空間ベースバーチャルスタジオの特徴の 1 つであるシームレスな合成機能を十分に発揮できる。

図 6.8 にバーチャルスタジオにおける各画像を示す。図 6.8(a)に示すように和室セットの一部分のみを残し、その他、実セットが無い部分は事前に撮影し作成された全天周画像より生成し(図 6.8(b))、カメラ映像(図 6.8(c))と合成する(図 6.8(d))。

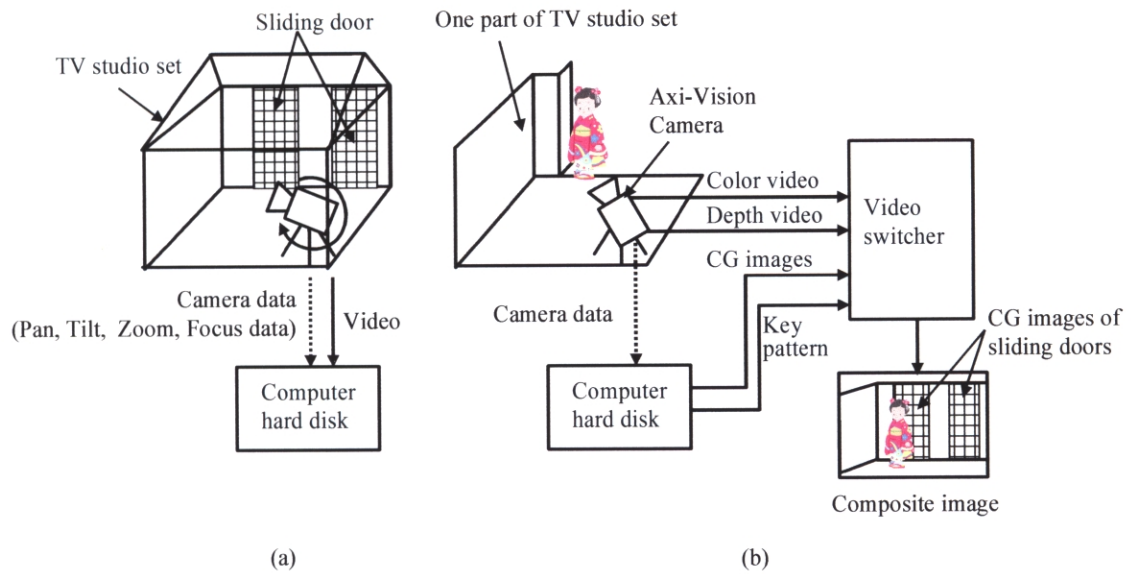


図 6.7 Axi-Vision カメラを使ったバーチャルスタジオ (a) 全周囲の映像を事前収録し、(b) 本番時には、カメラデータをもとに背景画像をリアルタイムで生成し、距離情報を用いて実写映像と合成する。

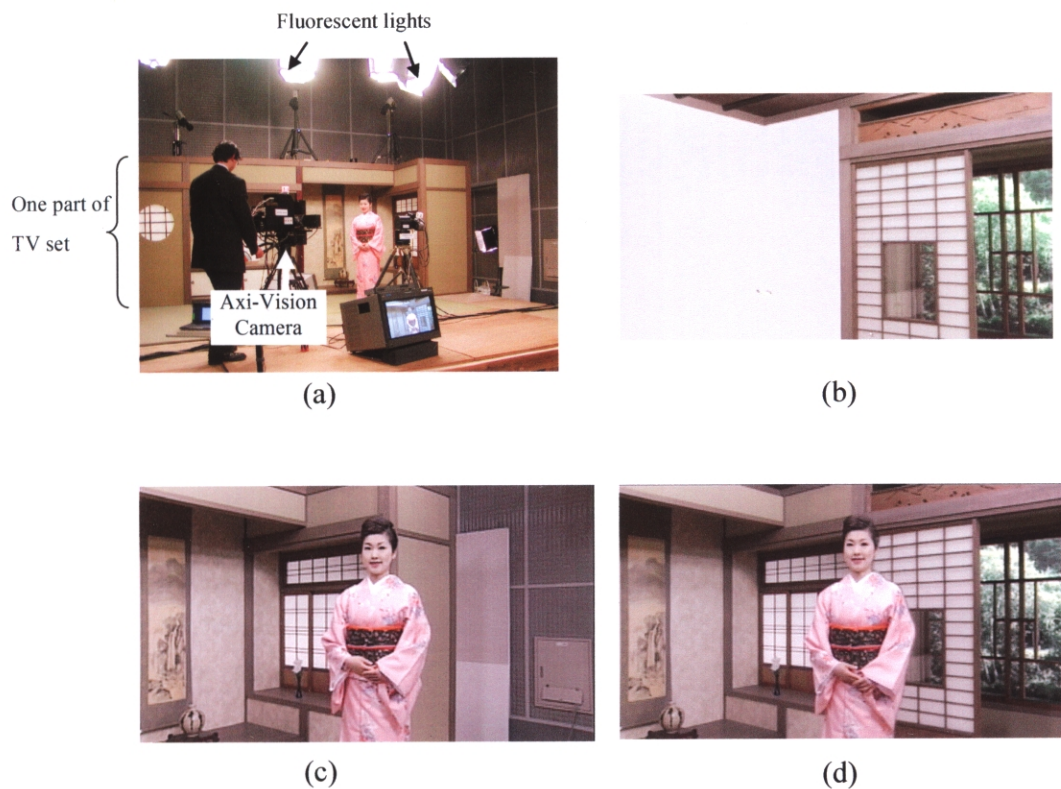


図 6.8 Axi-Vision カメラを使ったバーチャルスタジオの映像 (a) スタジオの様子 (b) CG から再生される合成素材映像 (c) 被写体の映像 (d) 合成映像

Axi-Vision カメラから実セットまでの距離は 3 m 以内、カメラからスタジオの壁までは 5 m 以上ある。図 6.8(c)のように背後に実セットが無い部分に出演者がいても、人物とスタジオの壁の距離関係をもとに、バーチャルセット(図 6.8(b))を自然に合成できる(図 6.8(d))。また、従来、特に大型スタジオセットの場合、ブルースクリーンの設置やそれを高輝度に照明する作業が大掛かりであったが、Axi-Vision カメラによりブルースクリーンが不要となることで、番組制作設備の効率化にも効果がある。

## (2) 番組応用

この Axi-Vision カメラを用いたバーチャルスタジオ(図 6.9(a))をハイビジョン番組の収録に応用した[11]。番組収録では、人物の手前や奥に CG の雪を降らせる合成(図 6.9(b))や人物の周囲を CG の文字が飛び回るシーン(図 6.9(c))などの新しい画像合成を実現した。Axi-Vision カメラでは、カメラレンズのズームも通常通り機能できるため、カメラレンズや雲台のデータとリンクした CG 合成が可能で、カメラワークに追従した自然な距離合成映像を演出できる。また、リアルタイム映像合成が可能であるため、収録時に出演者が CG との合成の様子を確認し、演技することができる。また、これまで類似の映像制作を行う場合には、映像の後処理作業に多くの時間とコストがかかるが、本手法では、完成に近い合成映像が撮影時に収録できるため、映像制作の作業量の低減にも効果があった。

さらに、本バーチャルスタジオ技術を生中継番組にも応用し、奥行き 10 m のスタジオでブルーバックなしのリアルタイム映像合成(図 6.10(a))や、実写と CG を 3 次的に融合した映像表現を実現し(図 6.10(b))、本カメラによるリアルタイムの距離合成技術の威力を発揮することができた[6]。

一方、課題として、出演者の毛髪や衣服の素材、出演者までの距離などの条件によっては、反射光量が少なく距離検出ノイズが多くなることがあげられる。この特別番組制作では、衣装や撮影シーンの距離検出特性を事前に調べるとともに、撮影シーンによっては距離画像をフィルター処理してノイズ補正した。Axi-Vision カメラ以外に複数のカメラによる生中継であったため、距離算出にともなうディレイタイムは、システム全体に遅延を入れ補償した。また、人物と足元の床面とはカメラからほぼ同じ距離であり、それらの画像を、距離情報をもとに分離することはできないため、この番組制作では床面への CG 合成は行わなかった。



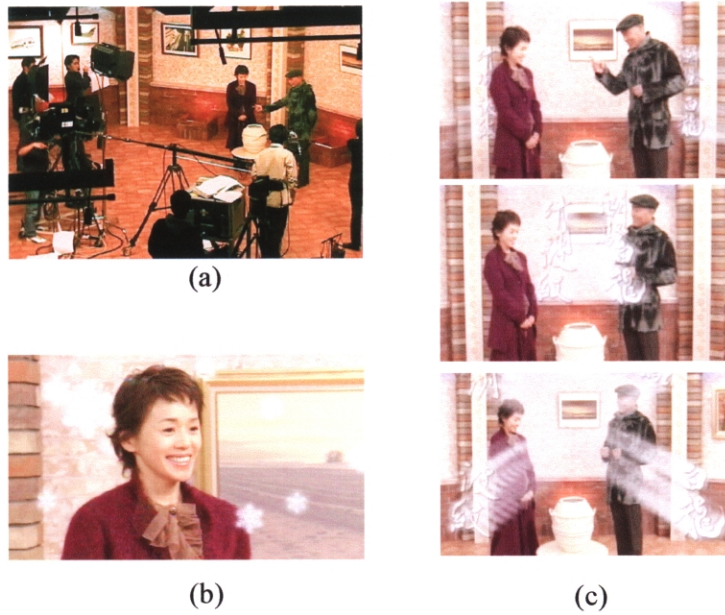


図 6.9 BS ハイビジョン番組「美しき日本 スペシャル」(2003 年 1 月 2 日放送)への番組応用 (a) バーチャルスタジオにおける番組収録風景 (b) CG の雪を人物の前後に合成した画像 (c) 文字を人物の周りに合成した画像



図 6.10 生中継番組「TV50 周年 今日はTVの誕生日」(2003 年 2 月 1 日放送)の映像 (a) 距離情報を用いた映像合成や (b) CG と実写の 3 次元映像合成をリアルタイムで実現した。

### (3) 定時放送番組への応用

2005 年 3 月 25 日から 9 月 25 日まで愛知県で開催された 2005 年日本国際博覧会“愛・地球博”会場に、NHK のハイビジョンスタジオ「グローバルスタジオ」(図 6.11)を整備し[12]、博覧会開催期間中、各パビリオンの紹介やイベント情報、視聴者からのメッセージ等を放送した。このスタジオに Axi-Vision カメラを配備し、距離情報を用いた映像合成技術を毎日の生放送番組の制作に効果的に活用した。このように定時番組に Axi-Vision カメラを使用するには、高品質な映像を長期間提供できる高い信頼性と、このカメラに関する専門知識のないオペレータでも容易かつ正確に扱える操作性が要求される。

このスタジオへの Axi-Vision カメラ応用に際して、距離検出特性の最適設定やノイズリデューサーの導入、CG 映像との合成手法の工夫により合成映像の質の向上を図った。まず、比較的広いスタジオ(面積  $120 \text{ m}^2$ )において、良好な距離検出特性を得るために、強度変調光の周波数を  $15 \text{ MHz}$  と低く設定し、第 3 章(3.8)式で表される測定レンジを約  $5 \text{ m}$  と長くした。また、強度変調光の位相調整により、検出精度が最もよい測定レンジ中心位置をカメラより  $3 \text{ m}$  の距離に設定し、出演者の立ち位置の目安とした。この強度変調光の周波数にあわせ、高速シャッター時間の半値幅も  $24 \text{ ns}$  まで広げ、反射光の受光量の増加を図った。さらに得られた距離画像に、リカーシブフィルター(Snell & Wilcox 社製, Niagra)による動き補正型ノイズ低減処理を加え、距離検出のノイズ成分を  $1/3$  に低減し、合成映像のエッジ部分の画質を改善した。

幾何学的オブジェクトや地球、地図などの 3 次元 CG は、画像処理装置(松下電器産業株式会社製、AV-VS3000、CG 生成能力 40000 ポリゴン/フィールド、テクスチャメモリ容量 1 GB)により、カメラデータに連動してリアルタイムに生成される。合成映像の生成法の一例を図 6.12(a)に示す。3 次元 CG を出演者の立ち位置で、前景と背景に分け、その間に Axi-Vision カメラの距離情報で抽出した映像を合成し、さらに最背面に Axi-Vision カメラのカラー映像を合成する(図 6.12(b), (c))。また、CG を半透明で合成する機能も活用するとより奥行き感や立体感のある演出効果が可能であるとともに、合成エッジ部分のノイズの影響も緩和される。

Axi-Vision カメラは、長さ  $20 \text{ m}$  のケーブルを介してスタジオ副調整室で一括制御され、カメラ起動時に初期設定が必要な CCD カメラや距離演算処理装置、信号発生器は、事前に調整した最適値にパソコンで自動設定するソフトウェアを開発し、安定運用を可能とした。また、番組制作時の調整箇所を I.I.の撮像ゲインと画像抽出閾値の 2 箇所とすることで、生放送にも迅速かつ容易に対応できる操作性を実現した。



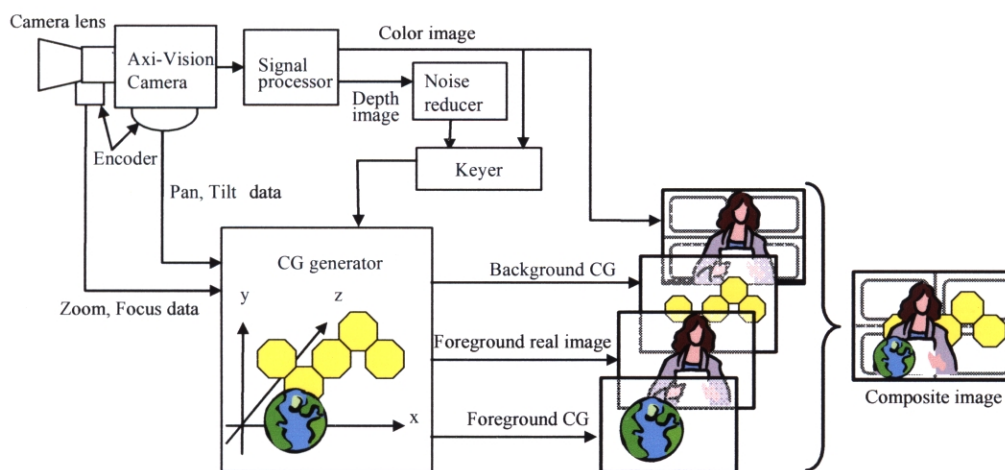
(a)



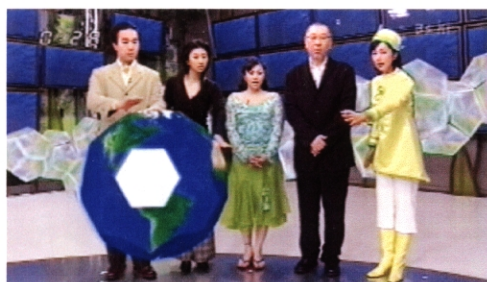
Axi-Vision Camera

(b)

図 6.11 2005 年日本国際博覧会“愛・地球博”特設の NHK グローバルスタジオ  
(a) スタジオ外観 (b) 番組制作の様子



(a)



(b)



(c)

図 6.12 距離情報を用いた CG 映像合成 (a) 合成方法 (b) 合成映像 NHK 総合、衛星ハイビジョン「愛・地球博がやって来る」(2005 年 2 月 26 日放送)  
(c) 合成映像, NHK 総合 東海地方ローカル放送「さらさらサラダ」(開催期間中 毎週月～金曜日放送)

スタジオの照明設備は、Axi-Vision カメラの使用を前提に蛍光灯照明で設計され、42W スポットライト 30 台と、300W フラッドライト 20 台でフロア全体をカバーしている。この蛍光灯照明は、消費電力が通常スタジオで使用されるハロゲン照明の約 1/7 であり、調光操作も可能となっている。また、スタジオ観覧用のガラス窓には、近赤外光カットフィルター（住友大阪セメント株式会社製、波長 850 nm の光の透過率 10.5%、590 nm の透過率 69.4%）を全面に貼り付け、近赤外波長の外乱光の影響を除去している。

スタジオフロアは、放送がない時間帯は周囲を取り囲む PDP 等による映像展示スペースとなるため、スタジオと展示スペースの切替えを頻繁に行う。このような場合、クロマキー合成法ではブルーバックの常設や出し入れ作業が伴い困難であるが、本カメラによる距離合成手法では、背面の PDP 等の有無にかかわらず、短時間で映像合成の環境が整えられ、機動的な番組制作が可能となっている。

スタジオでは、Axi-Vision カメラの他に通常のハイビジョンカメラ 2 台、ロボットカメラ 1 台の合計 4 台カメラが使用され、その場合、Axi-Vision カメラの映像系統とその他のカメラの映像系統の遅延時間の調整が課題となる。現在、Axi-Vision カメラでは、距離演算処理の遅延量が 1 ビデオフレームあり、さらに、CG 画像処理装置とノイズリデューサー内部で 3～4 ビデオフレームの遅延が加わる。対策として、すべての映像と音声系統に Axi-Vision カメラ系統と同じ遅延量を加える方法もあるが、このスタジオでは、機材の制限や生放送時の出演者の違和感にもつながるため、Axi-Vision カメラ映像の選択時のみ、音声系統に遅延を付加することで対処した。今後、距離演算処理と映像合成機能やノイズリデューサー機能を一体化したシステム構成とすることで、遅延量を減少できると考えられる。