

(続紙 1)

京都大学	博士 (情報学)	氏名	石 群
論文題目	Action History Volume for Spatiotemporal Editing of 3D Video in Multi-party Interaction Scenes (複数人物インタラクションシーンにおけるAction History Volumeを用いた3次元ビデオの時空間編集)		
(論文内容の要旨)			
<p>3次元ビデオは、3Dテレビのように映像が飛び出して見えるステレオ映像、もしくはコンピュータグラフィックスで作成された人工的な3Dアニメーションではなく、自由に運動する対象の3次元形状と動き、表面テクスチャを全周囲に渡って記録した完全な実写3次元映像である。3次元ビデオの編集には、従来の2次元映像に対する信号処理・画像処理的な編集だけでなく、撮影対象が持つ3次元的な形状および運動情報を考慮した時間的・空間的編集を行わなくてはならない。また、被写体が振袖や袴のような複雑な衣服をまとっている場合には、撮影された3次元ビデオから骨格構造を推定することが困難であるため、3Dアニメーションで用いられている骨格構造に基づいた映像編集法を適用することはできない。</p> <p>本論文は、個別に撮影された人物アクションの3次元ビデオを編集、合成し、複数人物によるインタラクションシーンを表す3次元ビデオを生成するための時間的・空間的編集アルゴリズムについてまとめたものであり、6章から構成される。</p> <p>第1章では、まず3次元ビデオの撮影プロセスを概観し、複数人物によるインタラクションシーンでは他者によって対象が隠ぺいされる現象が多発し、3次元ビデオを生成することが困難であることを指摘している。この問題は各被写体を個別に3次元ビデオ撮影し、これを編集・合成することによって解決されるが、新たな問題点として、個別撮影では被写体が互いの運動を参照できないために体の動きや視線方向に時間的・空間的な同期ずれが発生することを指摘し、この同期ずれを解消するための時間的・空間的編集としてAction History Volumeと名付けた新たな処理単位を用いた手法を提案している。</p> <p>第2章では、まず従来の3次元ビデオおよびコンピュータグラフィックスの編集手法について概観し、骨格構造を持たない3次元ビデオに対しては、撮影対象の3次元的な形状及び運動情報を考慮した編集が実現できないことを指摘している。次に従来の視線方向推定手法について概観し、自由に運動する人物の視線方向を非接触・非拘束な状態で計測する手法が確立していないことを指摘している。</p> <p>第3章では、まず、運動する対象の部位や視線がある時区間中に走査する3次元空間をAction History Volumeと定義し、Action History Volumeの表面領域が、その視線方向と部位・視線の運動方向との関係から、正(同方向)・負(逆方向)・零(その他)の3種類にラベル付けされることを導いている。続いて複数人物によるインタラクションの時間的・空間的編集が、Action History Volume同士の幾何学的な位置合わせ問題として定式化できること、さらに様々なアクション編集操作は3種類のラベルの可能な組み合わせである6種類のアクションパターンのいずれかによって表現され、各アクションパターンは個別に撮影された3次元ビデオの位置合わせ問題における制約条件として定式化されることを示している。</p> <p>第4章では、撮影対象の視線が定義するAction History Volumeを獲得するために、自由に運動する人物を撮影した3次元ビデオから、被写体の視線方向を推定する手法を提案している。提案手法のアイデアは被写体の顔を常に正面から撮影したかのような仮想正面顔画像を生成し、この画像から視線方向を推定する点にある。また仮想正面顔画像を高精度に生成する手法として、顔形状の左右対称性を制約として導入した高精度3次元顔形状推定手法と、この高精度顔形状を利用した超解像処理の導入を提案している。これらによって元の撮影画像以上の解像度で仮想正面顔画像を再構成することが可能であること、従来より正確な視線方向推定が達成されることを実際</p>			

のスタジオ環境での実験により示している。

第5章では、Action History Volumeを用いた複数人物によるインタラクションシーンの時間的・空間的編集アルゴリズムとして、(1)キー時区間、遷移時区間への分割、(2)キー時区間におけるアクションパターン制約を満たす解候補探索、(3)全時区間に渡る最適化という3段階からなる手法を提案している。まず(1)では、個別に撮影された各3次元ビデオを、インタラクションを伴うキー時区間とそれ以外の遷移時区間に分割するとともに、各キー時区間ペアに第3章で定義された6種類のアクションパターンうちのいずれの編集を行うか指示する。続いて(2)では、キー時区間ペア毎にAction History Volumeを作成し、アクションパターンとして与えられた制約を満たす編集パラメータを粗密探索によって探索し、解候補集合を求める。最後に(3)では、(2)で得られた各キー時区間に対する解候補集合から、元の動作からの変更量が最も少なく、運動の自然さを保存する解候補の組み合わせを動的計画法によって求めている。また提案手法を実際の3次元ビデオデータに適用することで、動作の自然さを損なうことなく、時間的・空間的編集が達成されていることを実証している。

第6章では、本論文の目的と提案手法のまとめを行うとともに、今後の課題と応用について議論をしている。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(続紙 2)

(論文審査の結果の要旨)

3次元ビデオは、3Dテレビのように映像が飛び出して見えるステレオ映像、もしくはコンピュータグラフィックスで作成された人工的な3Dアニメーションではなく、自由に運動する対象の3次元形状と動き、表面テクスチャを全周囲に渡って記録した完全な実写3次元映像である。3次元ビデオの編集には、従来の2次元映像に対する信号処理・画像処理的な編集だけでなく、撮影対象が持つ3次元的な形状および運動情報を考慮した時間的・空間的編集を行わなくてはならない。また、被写体が振袖や袴のような複雑な衣服をまとっている場合には、撮影された3次元ビデオから骨格構造を推定することが困難であるため、3Dアニメーションで用いられている骨格構造に基づいた映像編集法を適用することはできない。

本論文は、個別に撮影された人物アクションの3次元ビデオを編集、合成し、複数人物によるインタラクションシーンを表す3次元ビデオを生成するための時間的・空間的編集アルゴリズムについてまとめたものであり、得られた成果は以下の通りである。

(1) 運動する対象の部位や視線がある時区間中に走査する3次元空間をAction History Volumeと定義し、Action History Volumeを用いた身体運動の時間的・空間的表現法を考案した。

(2) Action History Volumeの表面領域が、その法線方向と身体の運動方向との関係から、正(同方向)・負(逆方向)・零(その他)の3種類にラベル付けされること、様々なアクション編集操作は3種類のラベルの可能な組み合わせである6種類のアクションパターンのいずれかによって表現され、各アクションパターンは個別に撮影された3次元ビデオの位置合わせ問題における制約条件として定式化されることを示した。

(3) 3次元ビデオとして撮影された被写体の視線が定義するAction History Volumeを得るために、自由に運動する人物の3次元ビデオから視線方向を高精度に推定するアルゴリズムを考案した。

(5) Action History Volumeを用いた複数人物によるインタラクションシーンの編集アルゴリズムとして、(1)キー時区間、遷移時区間への分割、(2)キー時区間におけるアクションパターン制約を満たす解候補探索、(3)全時区間に渡る最適化という3段階からなる手法を考案し、実写データによって有効性を示した。

以上本論文は、複数人物によるインタラクションシーンを対象として、骨格構造が未知の3次元ビデオの時間的・空間的編集という問題を解決するための方式として、Action History Volumeに基づいた時間的・空間的編集アルゴリズムを考案し、実際の3次元ビデオを用いた実験によってその有効性を実証したもので、学術上、実際上寄与するところが少なくない。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。

また、平成26年8月8日論文内容とそれに関連した口頭試問を行った結果合格と認めた。

注)論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。更に、試問の結果の要旨(例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」)を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。
要旨公開可能日： 年 月 日以降