

文化財記録ビデオ動画閲覧のためのメッシュスライダインターフェース Mesh-Slider Interface for Browsing Video Records

千田 智博†
Tomohiro Chida

川嶋 稔夫‡
Toshio Kawashima

1. まえがき

現存する文化財をデジタルアーカイブとして保存することは、文化財の保存のみならず、一般に公開するという観点で大変重要な役割を持つ。とりわけビデオカメラで動画撮影されたデジタルアーカイブは、静止画撮影に比べ、対象を様々な角度から網羅的かつ、連続的に撮影されることから、前後画像の位置関係や配置を把握しやすい。加えて、撮影者の撮影意図の変化をカメラの移動や時間経過で表現することが出来る。一方で文化財動画撮影では、クローズアップ映像を集中的に撮影したショットが大部分を占めるため、記録後に、いったいどの部分を視聴しているのか分からなくなるという点が問題になる。そのため、閲覧を支援、補助するシステムが必要になる。

現在、撮影された映像を用いて閲覧を支援する試みとして、多数の角度から撮影された画像を用いて三次元復元する手法[1]、複数画像や動画を用いてパノラマ画像や高解像度画像を生成する手法[2][3]などが提案されている。しかし、これらの手法は建造物や風景の全体像を撮影した映像に適した手法であり、文化財撮影のような特定部分を集中的に撮影した映像に適した手法ではない。また、カメラの動きや撮影角度に制限のある狭小空間においては、パノラマ化や三次元化に必要な撮影を行うことが出来ない。そこで我々は、特に、筒状や管状物体の内部構造における撮影映像の閲覧に有益な手法として、撮影対象と、実際に撮影された映像軌跡の概形を把握できるようなメッシュ構造を持つスライダ型ビデオインターフェースの設計を試みた。

2. 関連研究

2.1 折れ線スライダ

動画の閲覧を補助する試みとして、佐藤らは Coaster (COntinuous Access by Spatio-Temporal slidER) という概念を提唱している[4]。これは、映像の時間変化に伴う人物や物体の移動の軌跡を近似した折れ線をスライダとし、映像中へ配置するものである。例えば、被写体の動きに合わせて折れ線スライダを配置した場合、被写体の動きとマウスドラッグの動きが等しくなる。すなわち、被写体をつかんで移動させている様な可変速再生映像となる。

さらに佐藤らは、この技術を応用してパノラマビデオを提案している。パノラマビデオは物体の四方への移動が含まれる動画から、画像フレームの重なりを基にパノラマ画像を生成し、その中の移動物体の動きを折れ線スライダとする。そして、折れ線スライダ上をマウスドラッグした際の、マウスの位置に対応するフレーム画像を表示するもの

である。

パノラマビデオを実装した例として、佐藤らは走り幅跳びをする陸上選手を対象とした映像を題材として挙げている。幅跳びは助走から踏切までを横軸移動を行い、踏切後に頂点に達した後に着地を行う。これらの試技は、数秒程度で終えてしまうため、通常再生するだけの動画では観察に適していない。しかしこのパノラマビデオでは助走時の姿勢、踏切方、空中姿勢、着地姿勢など、各動作の再生部分が容易に把握でき、再生、停止等のビデオ操作も容易に出来る。このように Coaster を用いることで観察的、直感的操作でビデオ動画を閲覧することが出来る。

2.2 文化財記録ビデオの特殊性

パノラマビデオの概念をそのまま、文化財の一部分集中撮影映像、すなわち特定部分撮影映像に適応するには問題がある。パノラマビデオで扱われる映像は背景部分と移動物体に分けられ、背景は共通部分を用いて移動物体の軌跡であるスライダをその共通部分に配置することで、あたかも移動物体をつかんでいるかのように動画を再生するものである。しかし、今回対象にしているビデオ動画は一部分にクローズアップして撮影した、移動物体のない映像である。すなわち、パノラマビデオでいえば背景部分だけで構成された映像であることから、背景の一部分の移動軌跡を折れ線スライダにすることを考慮した場合、スライダ全体をそのまま画面に配置しても、マウスドラッグ操作と映像の移動量が同期する様な操作感を得ることはできない。

また、幅跳びの映像のように数秒程度の時間経過を表した映像ならば手でスライダの配置を行い、フレーム画像と軌跡の結び付けが可能であるが、複数本の数分の長さのショットを用いていた場合、スライダを手動で配置するという行為は時間的コストが大幅にとられることになる。そのため、折れ線スライダの生成、画面上への配置等の処理を自動化、あるいは半自動化し、処理時間を短縮する必要がある。さらに、複数のショットから作成した、それぞれの折れ線スライダを同じ平面上に合成する場合、同一撮影部分の関連付けや、カメラの向きを考慮した配置が必要になる。

これらの動画はカメラが入りにくい場所や、十分な作業スペースが確保できない状態を考慮し、また、既存のアーカイブ化されたビデオ動画に実用的な閲覧を支援するという目的を持つ。そのため、加速度センサや GPS といった特殊な撮影補助器具を用いなくて撮影されたビデオ動画を対象とし、動画内における映像の変化からシステム設計を行う。

3 メッシュスライダインターフェース

静物を対象として撮影を行う際は、ビデオカメラの注視点移動の軌跡自体を折れ線スライダとすれば良い。撮影者の自由なカメラ操作によって複数本のショットに同じ部分

† 公立はこだて未来大学大学院

‡ 公立はこだて未来大学

が写っていた場合、軌跡が交差するので、この交差部分を検出し、同一部分撮影映像として認識させることで、各動画間の位置関係を把握することが出来る。さらに、全動画を折れ線スライダとし、求めた位置関係を基に画面内にスライダを配置してゆくことで、網目状に交差するメッシュ型のスライダが完成する。この網目状のスライダは、交差部分を動画再生の切り替えポイントとし、現状見ている動画から、ユーザの操作に合わせて、見たい別のショットへ軌跡を乗り換えることが出来る。そのため、別ショットで撮られた動画でありながら、実空間上で近い距離にある部分を関連付け、集約的に再生することができる。以下にそのシステム実現のアプローチとしてメッシュ型スライダの生成モデル図(図1)と、本手法を導入したビデオブラウザの概念図(図2)を示す。

3.1 ショットからの折れ線スライダ生成

動画の動きをスライダへ対応づける方法として、オプティカルフローを用いる。取得したフローの中から、最頻ベクトルをフレーム間におけるカメラ移動方向として定めることが出来る。ベクトルの移動方向と移動量からアンカーポイントを定め、これらを結んだ折れ線が、折れ線スライダとなる。

3.2 動画間交差検出

複数の動画を撮影した場合、別ショットの動画で同一部分を撮影する場合がある。この重なった部分を同一部分であると自動的に判断し、網状スライダ生成時に交点部分とする処理を行う。そうすることで、交点では、折れ線スライダの進行方向切り替えポイントとして扱い、カメラ視点の分岐点として移動を行うことが出来る。

また、中空物体の内部構造撮影時などには、カメラを一回転させ、周回撮影を行う場合がある。この場合も、回転の開始点と終点を同一部分検出で見つけだすことで、動画がループして撮影されていることを知る事が出来る。図3に回転検出のモデルを示す一定方向にカメラが進んでゆき、元の場所に戻ってくるという特性から、オプティカルフローは同一方向に求められ、回転開始フレームと回転終了フレームは類似度が高くなる。

特徴点の検出方法には SURF(Speeded Up Robust Feature)を用いる。SURFは、物体認識に用いられるアルゴリズムであり、検出した特徴点の輝度勾配を求め、64次元の特徴ベクトルで特徴量を算出する。求められた特徴量をもとに類似度の高い特徴点同士を線で結び、結ばれた特徴点の数が閾値以上になった比較フレーム対を同一映像撮影部分とし、網状スライダの交点部分とする。

3.3 メッシュスライダの生成

複数本の動画を撮影した際に、ブラウザでは、動画内の移動量の軌跡に基づいて、折れ線スライダを実装することから、複数本の折れ線スライダが交わることになる。これを平面上に網目状にマッピングすることで、網状のスライダを生成する。このスライダは、記録された動画の進行方向とスライダの展開方向が一致しており、さらに交差する別動画のスライダへ自由に移行できることから、全動画像

内を直観的にカメラの移動を行えるインターフェースとして利用できる。

ビデオブラウザを実現するには、動画内の全フレームが適度にオーバーラップし、フレームの位置関係を交差検出によって相互に把握されていけばよい。実際の動画は、同一部分を自由な方向から撮影できるため、交差を持たないショットも存在する。そのため、目標とするショット内の映像と相互に交差のないショット同士の位置関係を知るとは困難である。そこで、自動でスライダを設定した後に、必要に応じて手動でフレームの位置を変更できる機能を付加する。

3.4 メッシュスライダの編集

メッシュスライダ生成後、類似画像のマッチングにおいて、異なる場所を撮影した複数動画の映像中に、似た特徴がみられると、その部分を同一部分であると誤認識し、本来、交差するべきでない部分で交差を作ってしまうことが起こりうる。また、カメラの手振れ等によって、オプティカルフローで推定した経路と、実空間上におけるカメラの移動経路に誤差が生まれることが起こりうる。

その際に、閲覧者が、メッシュスライダの編集を行うことで、より実空間に等しい映像空間を表現することが出来る。また、先述したように、この編集機能により、交差を持たないショットの動画を配置することが出来る。この場合、自動的に生成されたメッシュスライダの配置は候補として挙げられたものとして考え、利用者の操作によって再生、閲覧のし易いメッシュスライダを作成できる。

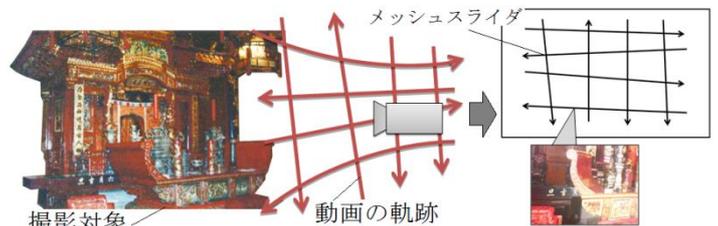


図1 メッシュスライダ作成モデル

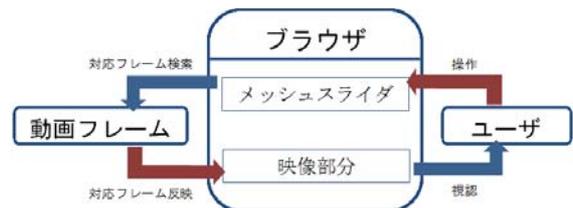


図2 システム概念図

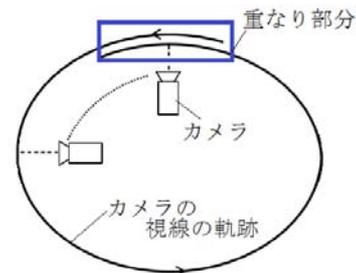


図3 カメラの回転撮影と重複の関係

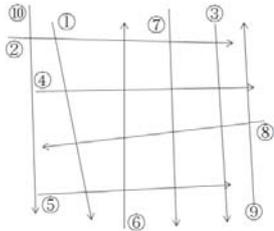


図4 各ショットの軌跡

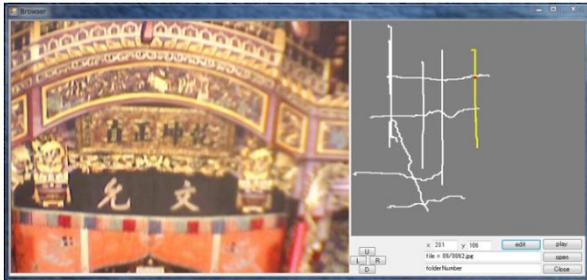


図5 システム画面のスクリーンショット

4. 実験

提案手法によって、実空間撮影時におけるカメラの軌跡と同型のメッシュ型スライダが生成されるか実験を行う。

次に作成されたメッシュスライダを編集することで、実空間に近い映像が展開されているかを検証し、得られた結果から本手法の有用性について述べる。

4.1 実験データ

実験に用いるデータセットとして、カメラ移動方向、交差する部分を調査済みのショット群を用いる。

撮影は室内で行い、照明は日常生活空間と同じ条件で行い、照明器具や三脚、および水平移動用のレールをはじめとする特殊な補助器具等はない。また、カメラが傾いているとオプティカルフローと、カメラの移動方向に差異が生まれてしまうため、カメラを大きく傾けないようにする。以上の条件で10秒から1分程度のショットを複数本用意する。

なお、各動画の軌跡は図4に示すとおりに撮影する。図中の数字は撮影した順番を、矢印は撮影方向を示している。

4.2 実験結果

この動画群を実装したシステムに適応した結果を図5に示す。同図における右上方の白色の網目構造が、動画の跡を射影したメッシュスライダで、黄色で表示されている軌跡が現在着目しているショット動画である。ユーザのドラッグ操作、もしくは右下方の操作パネルの入力により左方の映像部分に、操作結果に対応する動画フレームが再生される。

撮影に用いる動画の交差検出は類似画像のマッチングで行っていることから、カメラを固定してパン、チルト回転して撮影した動画や、カメラを壁面に対し平行移動させて撮影した動画でも同一のスライダ内に落とし込むことが出来る。

図4の実空間撮影動画の軌跡と図5右側のメッシュスライダを比較すると、求められた図が一致しない場合が見て

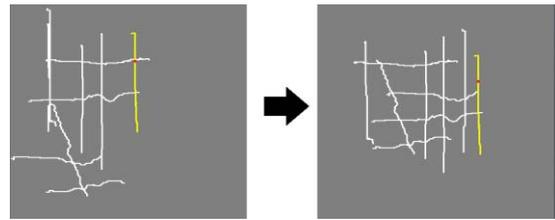


図6 スライダの編集(左:編集前, 右:編集後)

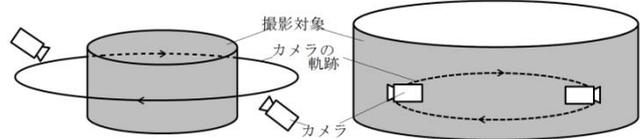


図7 カメラの一周回転映像のモデル

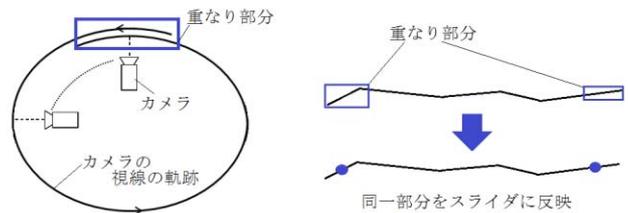


図8 スライダへの回転動作時重複部分の反映

取れる。具体的には、図4中の①と④、①と⑩、③と⑩のショット動画が挙げられる。これは、複数の折れ線スライダが近接し過ぎて、重なって表示されること、さらには、交差部分は類似検出で行っているため、基準となる折れ線スライダが実空間と異なる配置をとってしまうと、そのスライダと交差している、それ以降のスライダも配置が実空間のものとは異なることがある。そのため、閲覧者が表示されたメッシュスライダを編集する操作を行う。その編集結果が図6の右側画像である。

編集後に出力されたメッシュスライダ(図6右側)を図4の実際のカメラ映像の軌跡と比較してみると、実空間内の動画の移動軌跡が大まかな概形として射影されていることがわかる。

5. 課題

5.1 カメラの一周回転の検出

動画同士的位置関係を求める際に、交差検出だけでなく、同一動画内で回転が起こっているかを検出する必要がある。

カメラの回転動作にはいくつかのパターンがある。撮影に使用される回転は、図7左に示すパン・チルトからの回転、および、図7右に示す周回移動の回転の二つである。前者は中空物体の内部構造撮影時などに、カメラ位置を固定した状態で、一定の方向へパン、あるいはチルト回転を行い、360度カメラ視線を移動させるものである。後者は視線を撮影対象の中心部分へ向け、カメラはその周りを一周するように撮影するものであり、対象の全体を撮影する。

回転動作は、オプティカルフローを用いて軌跡を取得し、回転の開始点と終点部分が同一部分として認識することで検出できることから、カメラ動作と同様に、スライダをマッピングすることが出来る。

しかし、回転動作をスライダに反映する場合、オプティカルフローのみを頼りにマッピングすると、横、あるいは

縦に長いスライダとして反映されてしまう。この部分を解決し、同じ部分をループしながら参照しているということを閲覧者に明示的に知らせる必要がある。図8の例では、左側の重なりと認識された部分を一本のスライダとした場合、その部分をマーキングすることを表している。

5.2 同一ショット内の過剰交差

同一部分を多重に往復する場合、折れ線スライダが自身の動画フレームと交差する可能性がある。交差が多くなると類似する画像が一か所に集合することが出来るが、ユーザの求める画像がスライダから選択できない場合がある。これは、スライダが二次元平面上に展開されており、オプティカルフローやマッチングなどを基にマッピングされた画像の位置が、重複してしまい、スライダのクリック時に重複した画像のうち1枚しか表示できないためである。

6. 考察

今回実験に用いたデータから、出力されたメッシュ型スライダでは、実際のカメラワークと比較して、一致していない部分がある。その原因として、オプティカルフローの検出による移動方向の推定に誤差があるためと考えられる。以下にその原因として挙げられるものを挙げる。

6.1 照明条件に伴う移動方向検出

実験に用いた動画群では、照明の強く当たる部分と、物体の陰になっている暗所部分が存在する。暗所部分では、フレーム単位でみた場合、画面全体が黒一色になってしまう場合がある。このような場合はオプティカルフロー自体が検出されなかったり、実際の移動方向とは異なる方向へのフローが検出されてしまったりすることが起こりうる。これは、単色部分として認識される場合に起こる。すなわち、照明が当たっている部分でも色が一色であると、同様の結果が起こると考えられる。

それ以外の暗所を一部に含む場所の撮影では、フローが正確に求められていることから、表面形状や、構造物をもとに、エッジや画像の輝度の変化を認識できる映像であれば、検出精度に差は現れないと言える。しかし、カメラの性能によっては、絞りの自動調節速度、色相の再現度などが異なり、検出精度に差異が生まれてくる可能性がある。こうしたカメラのスペックによってはエッジや明度の移動量の検出が正しく行われない可能性がある。

6.2 表面模様、形状の条件に伴う移動方向検出

オプティカルフローの性質上、その検出には前後画像間の移動差異が明確に分かっている必要がある。

そのため、上述したような、単色で平らな面のみが表示されていてもフローは求めることが出来ないと考えられる。

また、表面の形状や、テクスチャによっては、オプティカルフローの推定結果がカメラの移動方向と一致しない場合がある。

例えば、通常のオプティカルフローは方向と移動量が一定の最頻値ベクトルが多く現れるが、細かい周期模様の場合は、方向と移動量が一定量の頻出ベクトルが複数存在する。また、進行方向を推定する際には、最頻出ベクトルを求めるが、頻出ベクトルが複数存在してしまうと、進行方

向とは異なるベクトルを最頻出ベクトルとして選択することが起こりうる。これがフローをバラバラに求めてしまう原因である。そのため、撮影の対象によってはオプティカルフローをカメラの移動方向推定の用途として使用するには難がある。

この問題は、映像内の大部分に細かい周期模様が存在する場合にも問題となる。例えば、細かいフローが多く検出されることで、移動方向のベクトルとは異なるベクトルが、最頻値ベクトルとして認識してしまう可能性がある。この問題を解決するには、ある程度の大きさを持つフロー以外を無視する処理などを加える必要がある。

また、動画のフレームレートにもよるが、近傍の表面模様や凹凸の間隔よりも比較フレーム間の移動量が少なければ、誤検出を防ぐことができると言える。そのため、細かく単調な模様や凹凸を撮影する場合は移動速度を抑える必要がある。

7. まとめ

本論文では、動画の時間的変化を空間的な変化に変換しユーザの空間把握を補助するインターフェースとその生成方法についての提案を行った。本提案の目的として、一部分の重点的撮影映像や狭空間撮影映像の閲覧支援を行なうことが挙げられる。狭範囲映像を記録したビデオ動画の閲覧システムとして、限られた動画の情報からそれぞれの動画の位置関係を推定し、動画の進行可能方向と現在の参照部分を表示する機能を持つ必要があると考え、現行の時間操作型動画スライダ(シークバー)に代わるインターフェースとして空間操作型のインターフェースとして折れ線スライダを採用した。

折れ線スライダの自動生成、さらに生成された折れ線スライダを結合させ、一つのメッシュ型スライダを生成する手法について実装例を挙げ、有用性をしめした。このシステムにより、フレーム間の時間的な距離の把握、動画の進行方向の事前把握、動画間の位置関係把握、実空間上の近辺画像フレームへの移動が行えるようになり、閲覧者がより観察記録者の立場になって文化財の閲覧を体験することを可能とした。より大規模な映像のアーカイブを取り扱うようにすることが今後の課題である。

8. 参考文献

- [1] 佐藤智和, “動画像からの三次元復元-マルチベースラインステレオ法を用いた屋外環境の三次元モデル化”, 日本工業出版, Apr.2003.
- [2] 千葉直樹, “画像特徴に基づく高速・高精度なパノラマ画像合成ソフトウェア”, Jun.2003
- [3] 池谷彰彦, 中島昇, 佐藤智和, “手持ちカメラによるビデオモザイクングおよび超解像”, 情報技術フォーラム(FIT)論文集, 2003
- [4] 佐藤隆, 阿久津明人, 南憲一, “Coaster:映像の時空間直観的操作による可変速再生方法とその応用”, 情報処理学会論文誌, Vol.40, No.2, pp.529-536, Feb.1999