

ズーム操作可能な多視点画像の閲覧方式

A Multiple View Images Browsing Method with Zooming Control

明智 那央† 北原 格† 大田 友一†
Nao Akechi Itaru Kitahara Yuichi Ohta

1. はじめに

映画マトリックスに採用されたことで広く知られることとなった Bullet-Time は、被写体を取り囲むように配置した複数台のカメラで撮影した画像を切り替えるカメラワークであり、近年は、映画だけでなく他の映像コンテンツにも採用されている[1][2]。多視点カメラの光軸が3次元空間中のある特定の点(注視点)で交わるようにカメラの位置姿勢を設定すれば、後は単純な画像の切り替え操作によって Bullet-Time 効果を実現することができる。しかし、カメラ設置のための時間を十分確保できなかつたり、撮影中に注視点が移動したりするなど、注視点を設定することが困難な状況が存在する。このような問題を解決するために、富山ら[3]は、2次元射影変換を用いて注視点を仮想的に一致させる手法を提案している。この処理では、ユーザが指示する注視点の3次元座標を(何らかの方法で)求め、その3次元座標が画像中心で観察されるような射影変換を計算し、撮影画像に適用する。

映像機器の発展に伴い、撮影画像の高解像度化が進んでいる。スマートフォンなどのモニタサイズの小さいデバイスを用いて、高解像度画像中の特定領域を詳細に観察する場合、対象領域をデジタルズームによって拡大して提示するのが一般的である。今後、撮影画像の高解像度化が進めば、Bullet-Time 鑑賞時にもこのような操作が起こりうると考えられる。図3(a)中の点線で囲まれた領域にデジタルズームインした状態で、多視点画像の切り替えを行う際、多視点画像上の同一箇所をクロッピングすると、図3(b)に示すように、視点の切り替わりに伴い注目物体が画面の外にはみ出してしまふ問題が存在する。この問題は、ユーザの注目点(ズーム中心)と多視点画像の注視点が異なるために発生する。

本論文では、この問題に着目し、ズーム操作可能な多視点画像の閲覧方式を提案する。ユーザの注目点と多視点画像の注視点が異なる場合には、ステレオ法を用いて注目点の3次元位置を推定し、その点を多視点画像の新しい注視点とすることにより、問題を解決する。

2. ズーム操作可能な多視点画像の閲覧方式

被写体を取り囲むように複数台のカメラを配置し、撮影した多視点画像に対する処理について述べる。ここでは、全てのカメラは事前にキャリブレーションされ、内部・外部パラメータは既知とする。ズームイン操作が行われた場合、[3]と同様の処理によって、ユーザの注目点と多視点画像の注視点を一致させる射影変換行列を求める。その射影変換行列にズーム操作によるスケール変換行列を積算する

†筑波大学, University of Tsukuba

ことにより、図3(c)に示すような、ズーム操作可能な多視点画像切り替えを実現する。

2.1. 注目点の3次元座標推定

まず、ユーザの注目点が画像の中心となるような射影変換を行うために、注目点の3次元座標を推定する処理について、図1を用いて説明する。ユーザは、画像I0を見ながら、ズームインする領域(注目点:図1の緑点)の2次元座標 $Z0(u0, v0)$ を、画面上のポインティング操作などを用いて与える。画像I0とは異なる視点で撮影した画像I1との間の基礎行列 F をカメラパラメータから算出し、図1の青線で示す $Z0$ に対応するエピポーラ線を求める。エピポーラ線において、 $Z0$ との類似度(本システムでは SSD : Sum of Squared Difference) を計算して対応点探索を行うことで、図1の赤点に示す対応点を得る。対応点情報に対してステレオ視を適用することにより、ユーザの注目点の3次元座標を推定する。

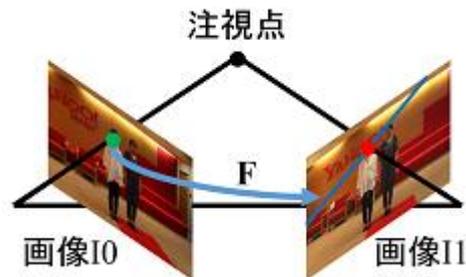


図1. 注目点の3次元座標推定

2.2. 射影変換処理

推定した注目点の3次元座標が画像中心で観察されるように、各多視点画像に射影変換を施す。各カメラの光学中心を原点とし、注視点方向を Z 軸とする直交座標系を考え、各軸の単位ベクトルを ex, ey, ez とすると、前節で求めた注目点を画像中心に合わせるための回転行列 R'_m は次式となる。この射影変換処理により注目点と注視点が一致する。

$$R'_m = (ex \ ey \ ez)^T$$

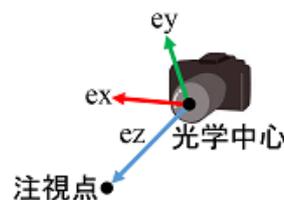


図2. 各カメラの新しい直交座標系

Bullet-Time で生成させるビデオシーケンスの連続性を維持するためには、注目物体の見た目の大きさをほぼ一定に保つ必要がある。本システムでは、提示画像を撮影する仮想カメラの内部パラメータ行列 A'_m を次式のように定義する。ここで、 f'_m は各カメラの焦点距離の平均値 f_{ave} に、ズームイン操作を行ったカメラから注視点までの画像距離 w_i と各カメラの注視点までの距離 w_m との比を掛け合わせた値とする。 C_x と C_y はそれぞれ画像中心の x 座標と y 座標である。

$$A'_m = \begin{bmatrix} f'_m & 0 & C_x \\ 0 & f'_m & C_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$f'_m = f_{ave} \frac{w_m}{w_i}$$

さらに、撮影時のある多視点カメラの姿勢を表す回転行列を R_m 、そのカメラの内部パラメータを A_m 、スケール変換行列を S とすると、これらの合成行列として、2次元射影変換行列 H_m が求まる。

$$H_m = SA'_m R'_m R_m^{-1} A_m^{-1}$$

最後に、求めた2次元射影変換行列 H_0 で画像 I_0 を射影変換することにより、 Z_0 を注視点にし、かつ、その点にズームインした画像 I_{Z_0} を得る。他の視点の画像に対しても同様の処理を施すことにより、ズームインした多視点画像を得る。これらを切り替えながら鑑賞することで、Bullet-Time 効果を実現する。

3. 実装結果

2章で述べた処理を実装することで、ズームイン操作可能な Bullet-Time ビューアを実現できる。しかしながら、本研究の対象は高解像度画像であるため、CPU で射影変換処理を行うと計算コストが高く、リアルタイムに処理することができない。そこで、GPU でピクセルごとに並列で射影変換処理を行うことで、リアルタイム処理を実現する（本システムでは GLSL を用いて実装を行った）。

実行環境は CPU が Intel Core i5 1.70GHz、GPU が NVIDIA GeForce GT 620M、メモリは 10.0GB RAM のノート PC である。閲覧対象の画像の解像度は 5616 画素×3744 画素で、提示画像はその 8 分の 1 の 702 画素×468 画素とする。以上の構成で構築したシステムを用いて実証実験を行い、リアルタイム実行可能であることを確認した。図3に提案手法による処理の一例を示す。図3(a)は撮影した多視点画像の一枚であり、緑点が撮影時に設定した多視点画像の注視点である。ユーザによって入力された青点を注目点として、白色の破線で囲まれた領域にズームインした場合を考える。図3(b)に示すように、多視点画像の同じ領域を単純に切り替えると注目点と注視点が異なるために、視点の切り替わりに伴い注目物体（この例では顔）が画面外にはみ出してしまう。提案システムを用いてズームインしながら多視点画像の切り替えを行った結果を図3(c)に示す。視点が切り替わっても、注目物体が常に画像の中心で観測されている。

4. おわりに

ズームイン操作可能な多視点画像の閲覧方式を提案した。従来の単純な画像切り替えでは、注目点と多視点画像の注視点が異なる場合、視点移動に伴い注目対象が画面外にはみ出してしまう問題があった。ズームイン操作の際に、注目点を多視点画像の注視点に設定し直すことでこの問題を解決した。射影変換処理を GPU で並列計算することで、リアルタイム処理可能なシステムを構築した。今後の課題としては、注目点の3次元座標推定の精度の向上や、タブレットやスマートフォンでの実装を検討している。

参考文献

- [1] <http://www.ri.cmu.edu/events/sb35/tksuperbowl.html>
- [2] <http://www.timeslicefilms.com/>
- [3] 富山仁博, 宮川勲, 岩館祐一, "多視点ハイビジョン映像生成システムの試作: 全日本体操選手権での中継番組利用", 電子情報通信学会技術研究報告, PRMU, パターン認識点メディア理解, Vol. 106, No. 429, pp. 43-48 (2006).



(a)



(b)



(c)

図3. 単純なズームイン画像と提案システムによるズームイン画像の比較