

映像メディアに対する 時間的および空間的インタラクション手法の提案

高嶋 章雄[†], 山本 恭裕[‡], 中小路 久美代[§]

概要: 本研究は、映像メディアを時間的変化を伴う可視化表現 (TbVP: Time-based Visual Presentation) としてみなし、その表現とのインタラクションを通じて、メディアの内容の理解を支援することを目指すものである。TbVP の表示速度を変化させることによってユーザの得る情報の種類や質が大きく異なることを利用し、様々な表示速度で映像の空間的な任意の部分を表示させる枠組みを構築した。本論では、表示速度を変更する時間的なインタラクションと、表示される見たい目を変更する空間的なインタラクションを可能とする環境のためのモデルおよびプロトタイプシステムについて説明を行う。

Spatial and Temporal Interactions with Video Data

Akio TAKASHIMA^{††}, Yasuhiro YAMAMOTO^{‡‡}, Kumiyo NAKAKOJI^{§§}

Abstract: The goal of this research is to support analyzing of video data as Time-based Visual Presentation (TbVP). Changing the browsing speed of a TbVP gives a user the qualities of information. We have developed a framework where any part of frame can be displayed with different browsing speeds. This paper first describes a model for an environment, which allows a user to have spatial and temporal interactions with video data, and then explains prototyped systems based on the model.

1 はじめに

ビデオやアニメーションなどの映像メディアを計算機上で再生、閲覧することが可能となってきた。このような映像メディアの内容には様々な種類があり、それに応じて適切な閲覧方法も異なってくる。たとえば映画や CG アートなどの映像作品は、それを作る側がどのように見せるかまで考慮して作成した映像メディアである。したがってこのような映像を閲覧する際には作者が想定した見方で映像を見るのが望ましい。

一方で、実験ビデオや、動的なシミュレーション結

果をアニメーションとして再現するものなど、ある事象の分析を目的として作成された映像においては、事象の持つ様々な情報を見る側がより多く知ること、または見る側が必要とする特定の情報を知り、映像の内容を理解することが閲覧の目的となる。たとえば野球の試合をビデオ撮影し分析することを想定すると、単に勝敗を知りたい場合や、スコアも必要な場合、あるいは選手のバッティングフォームを知りたい場合など、ユーザが得ようとする情報は様々である。このような状況においては、ユーザが分析を進めるうちに必要とする情報が明確になる場合もしばしばある。こういった探索的分析を目的として映像を閲覧する際には、前述したような受動的な映像の見方ではなく、ユーザが能動的に映像を見る必要がある[1]。特に、映像に対してユーザが時間的、および空間的にインタラクションを行うことが、その内容の理解に役立つ[2]。

[†]奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology

[‡]科学技術振興事業団 戦略的創造研究推進事業(さががけ研究 21)
PRESTO, JST

[§]東京大学先端科学技術研究センター
Research Center for Advanced Science and Technology, University of Tokyo

本研究は、映像の内容をユーザが理解するために探索的な分析を行うことを想定し、ユーザが映像メディアと自由にインタラクションを行える環境を提供するものである。

本稿では、まず映像内容を理解するための表示速度の変化について述べ、時間的、および空間的なインタラクションを可能にする TbVP (Time-based Visual Presentation) の概念を説明する。次いで、その概念に基づき作成したプロトタイプシステムについて説明する。なお映像メディアとは、一般的には画像と音声などの情報が同期して出現するデータストリームであるとされる場合が多いが、本稿においては、動画や映像といった視覚的なデータストリームのみを指すものとする。

2 映像内容理解のために

本章では、ユーザが映像の内容を理解するためには、どのような時間的、空間的表現が効果的であるかについて述べる。

映像を閲覧する際に、その表示速度を変化することで、ユーザが得る情報の種類や質が大きく異なる。たとえば、サッカーの試合のビデオを観察する際、選手のポジショニングの変化やチーム全体のフォーメーションの変化などを理解するには、ある程度以上の速い表示速度で映像を見て流れを掴む必要がある。一方でボールにかけられた回転や、選手のシュートフォームなどを詳しく理解するには遅い表示速度で観察しなければ理解することが難しい。つまり表示速度を様々に変更した映像をユーザに提示すれば、映像メディアからより多くの情報を得ることができ、内容を理解することに繋がると考えられる。

表示速度を変更することは、換言すると映像メディアの持つ時間を、どのようにユーザに経験させるかを変更するものである。映像メディアの持つ時間とは、映像メディアデータを既存のムービープレイヤーを用いて1倍速で再生した際、最初から最後まで見終わるのに必要となる時間を指す。たとえば90分のサッカーの試合全部を録画したホームビデオや、台風が1週間かけて大陸を縦断する様子を10秒の映像にしたものなどでは、90分、および10秒が映像メディアの持つ時間長となる。本研究ではこれをMDT (Media Data Time) とよぶ。それに対し、映像メ

ディアを見ているユーザの置かれている環境の時間(実時間)は常に流れ続けており、映像を再生している途中で一時停止した場合でもユーザは常に流れ続ける時間を経験している。このような時間を本研究ではUET (User Experience Time) と呼ぶ。本研究は、MDTをUETに変換してユーザに経験させることで、映像を様々な速度表現として見せ、それによりユーザが様々な情報を得ることができることを目指すものである。

さらに、前章で述べたように、探索的な分析においてはユーザが受動的に映像を見るのではなく、能動的に様々な見方を試し、何らかの情報を理解しようとするプロセスが必要である。そのための映像に対するインタラクションを提供することで、ユーザが試行錯誤を繰り返しながら映像の内容を理解していくことが可能となる。MDTからUETへの変換についても、ユーザが映像を見ながら自由にインタラクションを行い表示速度を変えることで、ユーザの得る情報も様々に変化し映像の内容の理解が進むと考えられる。

映像に対する時間的なインタラクションを行う関連研究としては、佐藤らが映像の表示スピードをインタラクティブに変更するものとして折れ線スライダを提案している[3]。これは、映像中の物体などが移動する軌跡に沿って時間をコントロールするスライダを配置し、物体をドラッグすることで映像の可変速再生を可能にするものであるが、あらかじめ物体の動きを把握しておく必要があり、映像内容を理解した次の段階で利用可能な手法である。また、Andrienkoらの研究[4]は、映像中のオブジェクトの動き方を理解する探索的なタスクを支援するものであるが、主としてオブジェクトの軌跡の表現をインタラクティブに変化させて動きの変化量を見やすく表すものであり、動く様子そのものを積極的に利用するものではない。Liらの研究[5]では表示速度をインタラクティブに変更するシステムを作成しているが、少ない時間で効率良く映像を見ることに注力されており、探索的な映像の分析を目的とするものではない。

映像の内容を理解するために時間の流れを変化させて提示することについて述べたが、実際に内容を理解するには視覚的な表現が非常に重要な要素となる。たとえば時間の流れ方の異なる複数の映像

を比較しながら内容を理解する場合や、映像の空間的な一部分にのみ着目したい場合など、映像の視覚的な表現方法も様々に変更しながら見ることで内容の理解につながる情報を得ることができる。よって時間的なインタラクションとともに、このような見た目の空間的なインタラクション、すなわちインタラクティブな可視化表現を考慮することも必要となる。

そこで本研究では、映像の表示速度を変更する時間的なインタラクションと、表示される見た目を変更する空間的なインタラクションを可能とする可視化表現として、TbVP(Time-based Visual Presentation)を提案し、実際にインタラクションを行う環境としてTbVP Browserを構築している。次章ではこれらについて詳述する。

3 TbVP(Time-based Visual Presentation)

本章ではTbVPの説明とそこで可能になる時間的、および空間的インタラクションについて述べる。

3.1 TbVP利用の枠組み

TbVPとは時間的変化を持つデータに対して、その時間の流れ方と計算機上での見せ方をユーザが自由にコントロールできる可視化表現のことを指し、ディスプレイ上では映像または動画像として表示される。図1に示すようにTbVPは時間的変化を持つデータに対して以下の2種類の変換を行うことにより作成される。

- Time-based Transformation
- Visual Transformation

Time-based Transformationは、前章で説明したMDTをUETに変換するもので、Visual Transformationはユーザが見る視覚的な表現を変更するものである。TbVPが作成された後も、ユーザはTbVPを見ながらそれぞれのtransformationを変更することができ、ディスプレイ上の可視化表現をインタラクティブに変更することができる。

ユーザがそれぞれのtransformationを利用して時間的および空間的インタラクションを行うための環境として、TbVP Browserを構築した(図2)。TbVP Browserでは、ひとつの時間変化を持つデータから複数のTbVPを作成し、それらを比較することのできるインタラクションを提供する。さらに複数のTbVP間

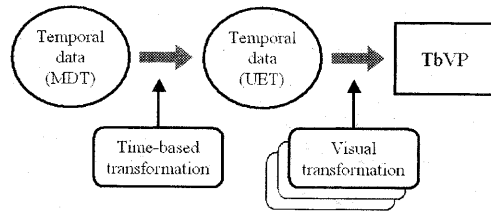


図 1: TbVPの作成

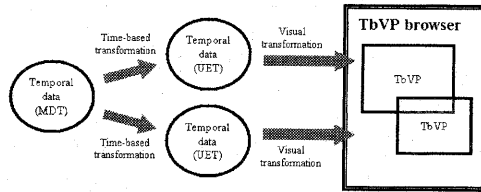


図 2: TbVP Browser

で時間的、空間的な同期をとることも可能としており、表現されている複数のTbVP間の関連性を保つことができる。このようにユーザがTbVP Browser上でTbVPに対するインタラクションを行い、試行錯誤しながら映像の理解を深めていくことが期待できる。

以下に、各transformationの説明、およびそれらを利用したTbVPとの時間的、空間的インタラクションについて述べる。

3.2 時間的インタラクション

時間的インタラクションとしては大きく以下の2つに分けられる。

- 既存のムービープレイヤーでの時間制御
- rateの変更

それぞれについて以下に説明する。

Time-based TransformationによりMDTをどのようにUETに変換するか、つまりどのように表示速度を変化させていくかはrateの変化により実現される。rateとは、映像の表示速度(speed)と映像が再生される方向(direction)からなる。

一般的なムービープレイヤーにおいては、再生や停止、早送りボタン、タイムスライダ[6]など時間の流れを制御するコントローラがついている。TbVPの枠組みに当てはめると、これらもMDTをUETに変換するものとなる。たとえば、再生は、MDTとUETを1対1に対応させ、rate1倍の変換で見ることである。

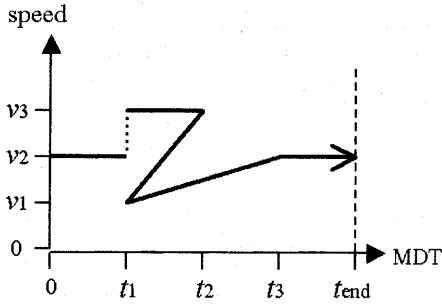


図 3: rate 変化グラフ

同様に、停止は rate が 0 倍で UET がどれだけ経過しても MDT が変化しない状態を示し、3 倍速の早送りは rate が 3 倍の状態、MDT を本来の 3 倍の早さで UET として経験することとなる。また、タイムスライダは、MDT のどの時点の画像をディスプレイ上に表示するかを決定するものである。これらは MDT を UET に変換する非常に単純な例であり、主として映像メディアを受動的に見る、あるいは表示内容を特定の時点へ移動する際に利用することができるが、探索的な映像の分析プロセスを支援するには不十分である。

本研究ではこれらのインタラクションに加えて、Time-based Transformation における表示速度の変化をユーザがインタラクティブに行えるようにする。1 章で述べたように、映像の内容理解のため情報を様々な視点から読み取ることが可能にするには、より複雑な表示速度の変更を可能にする必要がある。そこで本研究では速度変化だけでなく再生する方向も制御できるよう前述の rate を利用し、rate の変化を図 3 に示した矢印付きのグラフで表現する。図 3 では、まず MDT の最初(0)から t_1 までにあたるデータを v_2 の速度で順方向に表示し、 t_1 を過ぎた瞬間から t_2 まで速度 v_3 で表示する。その後、表示速度を v_3 から v_1 まで遅くしながら t_1 まで逆方向に表示する。その後 t_3 まで速度を上げながら順方向に表示し、 t_3 へ達すると v_2 の速度で最後 (t_{end}) まで表示する。グラフが示す速度で MDT の最初から最後まで矢印をなぞることで UET、つまり実際にユーザがどのような時間の流れで TbVP を見るかがわかる。

ユーザが自由に表示速度とのインタラクションを行えることが重要であるが、この rate 変化のためのイン

タラクションとしては、まず、実際にグラフ記述を操作することが考えられる。表示速度が様々に変化する状態を、TbVP を見ることだけで認識するのは困難であるため、それをどのように変化させているかという感覚をユーザが失わないようにするために有用である。rate を操作するためのインタフェースと、現在操作している状況を見るためのインタフェースを兼ねており、実際には図 5 の下段のような実装となっている。しかし、実際に TbVP を見ながらその内容を理解しようとする際には、グラフを記述するというインタラクションが効果的であるとは言い難い。そのため、ある程度ユーザにとって意味のある rate 変化のパターンを用意しておきそれらに特化したインタラクションを提供することも考慮する必要がある。詳細は次章で述べる。

TbVP Browser 上で複数の TbVP を表示している際にはそれらの時間的な同期をとることができる。時間的な同期とは、rate の変化により表示している MDT の時点にずれが生じている複数の TbVP を、同じ MDT、同じ rate に合わせることを指す。複数の TbVP を同期させることの利点としては、一連の rate の変化を、映像を見ながら何度もインタラクティブに繰り返すことが可能となることや、複数の時間の流れで表現される TbVP の関連を明確にすることなどが挙げられる。

図 4 は可変 rate と固定 rate で表された 2 種類の TbVP に関して、UET に対するそれらの MDT の進み具合が、等しい状態から一旦異なり、再び同じ状態に同期する様子を表している。図 4(a) は MDT に対する rate の変化を、(b) はその変化に基づいた MDT から UET への変換を表しており、どちらの図も点線が一定の rate を保ったもの、実線が変化する rate を表している。固定 rate の UET は時間の経過に関らず一定の速度で順方向に MDT の示す映像を表示する。それに対し可変 rate の UET は target (ユーザが興味を持った時点) となる MDT の特定の時点での映像が表示された直後、一定の時区間を高速で逆方向に表示し、その後 target に近づくにつれて徐々に表示速度が遅くなり、target 付近で一定の速度を保った後、点線で示された MDT に同期するまで速度を速くするものである。(b) では UET 軸の上から下に向かってユーザが 2 種類の TbVP を見ている事と

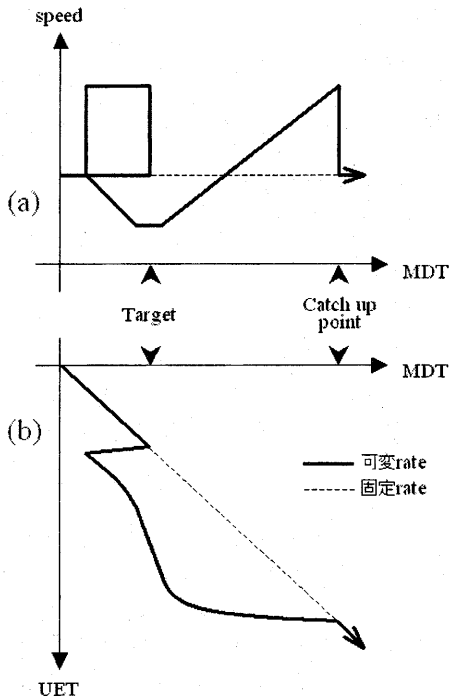


図 4: 時間同期

なり、同じMDTの時点の映像を表示していた2つのTbVPが、一度異なる時間を表現し、再び同期することがわかる。

ここでは、可変rateが固定rateに同期するまで速度を速める例として述べたが、同期するcatch up pointを指定することや、同期までのrateを指定することでも実現することが可能であり、ユーザーが自由に選択できる方が望ましい。

3.3 空間的インタラクション

一般的なムービープレイヤーにおいては、視覚的、空間的操作のためのインタラクションを提供しているものは少なく、ディスプレイ上でのサイズを変更できる程度である。ビデオ編集ソフトではビジュアルエフェクトを用いて見た目を変化させることができるが、インタラクティブなものではない。映像を新たに作成することが目的であるため、映像の理解といった本論で対象としているようなタスクのために用いるには不向きである。

本研究で用いた空間的インタラクションは以下の4つであり、いくつかを組み合わせる利用することが

できる。

TbVP 単体に対するインタラクション

- 表示位置の変更
- 透明度の変更

複数の TbVP を用いるインタラクション

- 全領域の同期
- 部分領域の同期

本研究における TbVP に対する空間的インタラクションは、主として TbVP Browser 上に複数の TbVP が表示されている際に有用である。たとえば、時間の流れ方の異なる複数の TbVP を同時に表示している状態で、TbVP Browser 上の空間的な配置をユーザーが変更することで比較を容易にすることや、それぞれの TbVP の透明度をユーザーが任意に変化させることで、どの時間の流れに着目するかをインタラクティブに変化させて観察するなど、様々な見方が可能となる。

時間同期と同様に空間同期も利用して、ユーザーが映像を理解するための様々な見方を提供することができる。空間同期とは、TbVP の表示される空間が特定されていることを指す。図 6 では人間が走る映像を例に2種類の空間同期方法を示している。図 6 の最上段に書かれたフィルムのイラストは、人間が走る映像の例を紙面で表現するために用いたものである。フィルムの左が早い時点、右が遅い時点を表しており、人間は映像内の空間を左から右に向かって走っている。図中の TbVP(1) はフィルムの左端の映像を、TbVP(2) は左から4番目の映像を表示しているものとする。

図 6 (a) は表示する空間を定めずに、時間の流れ方が異なる2つの TbVP を異なる空間に配置した非同期の状態である。

(b) がそれらを正確に重ね合わせたもので空間同期の取れた状態である。この状態でそれぞれの透明度を変化させることで、たとえば速い再生速度の TbVP と遅い再生速度の TbVP を同じ空間で比較することができる。あるいは表示速度の等しい TbVP に時間的なずれを持たせて表示することで、残像を残しながら人間が走るような表現として観察することもできる。これらはカメラワークのない固定された視点で取られた映像の動きを比較することなどに利用することができる。

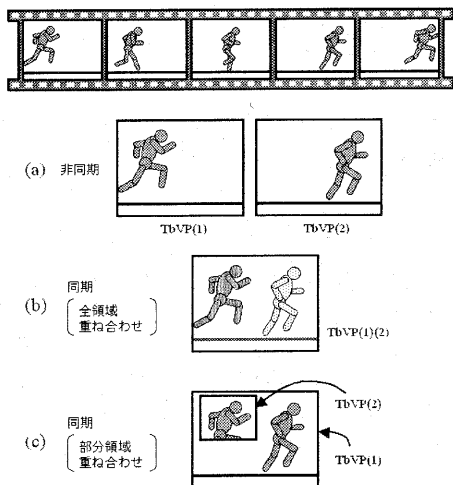


図 6: 空間的同期

(c)は TbVP(2)の領域を小さくし、TbVP(1)と部分的に同期させた状態を示している。実際には異なる時間の流れを持つ2つのTbVPが表示されているが、TbVP(1)の空間的な一部分の時間の流れを変更しているような表現として見る事ができる。図では、TbVP(1)で既に過ぎ去った人間のいた場所における、腕の振りをユーザが観察している状態である。他にもサッカーの試合におけるゴールの位置など、映像の内容に空間的な意味がある場合に有用である。

4 プロトタイプシステム

前章で述べたようなインタラクションを可能にする TbVP Browser のプロトタイプシステムを構築した。

4.1 任意のrate変化

図 5は時間の流れを操作する Rate Controller を持つプロトタイプであり、複雑な rate 変化をユーザが作り出すことができる。Generic Controller を利用することで、ユーザは表示されている全ての TbVP に対し、再生と停止、頭出し、および Presentation Slider による MDT の表示時点の設定が可能であり、既存のムービープレイヤーと同等の機能を持っている。

図では 2 種類の TbVP が表示されており、そのうち領域の大きいものは 1 倍速での表示(既存のムービープレイヤーによる再生と等速)であるが、領域の小さい TbVP は、画面下部の Rate Controller によ

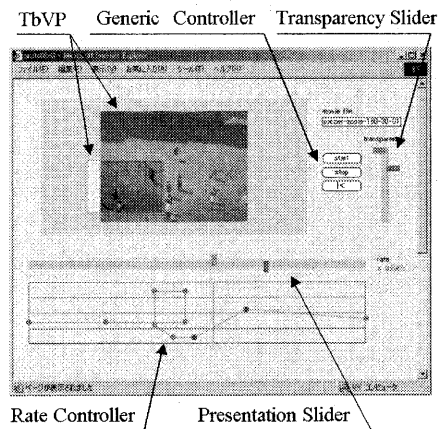


図 5: スクリーンショット

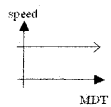
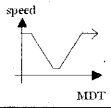
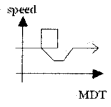
て表示速度を制御されている。Rate Controller は図 3に示すものと同等のもので、横軸が MDT、縦軸が rate の速度を表している。左端が MDT の最初を、右端が最後を表しており、TbVP が順方向に再生されるにつれて Presentation Slider のインジケータが左から右に動きその位置に対応する MDT での映像が表示される。またその際のインジケータの移動速度、および TbVP の表示速度がグラフの示す rate の速度となっている。マウス操作によりグラフの形を調整することで、ユーザは MDT をどのように経験するかを決定する。

ここで表示されている 2 つの TbVP は空間的に同期している状態である(部分領域の同期)。片方の TbVP のサイズや位置変更などのマウスを用いた空間的なインタラクションにより、表示されている映像中の興味を持った空間の一部分を選択的に観察することができる。たとえば選手がボールを蹴る瞬間を時間と空間で指定して詳しく見ることが可能になる。また、Transparency Slider によりそれぞれの映像の透明度を変更することで、ユーザが着目したい映像を見やすく表示することも情報の理解に役立つと考えられる。

4.2 意味のあるrate変化パターン

図 5では rate controller による rate 変化を自由に行うことが可能であるが、前述したように実際に映像を見ながら内容を理解しようとしている際、マウスでグラフを形成する操作はあまり直感的であるとはい

表 1: rate 変化のパターン

	想定するタスクシナリオ	Rate 変化のパターン	特徴	必要となるコントローラ
Type1	全体を速い速度で見て、流れを掴む。 全体を遅い速度で見て、細かな動きを観察する		常に一定のrateを保つ	保持するrateを設定する
Type2	興味のある部分をゆっくりと観察し、それ以外の部分は流し見をする		targetに近づくほど遅いrateで表示し、遠ざかるほど速いrateで表示する	MDTのどの時点(部分)をtargetとするかを決定する
Type3	観察中に気になった瞬間をもう一度確認する		ユーザがtargetを決定した直後に、MDTを現在の時点より巻き戻して、target付近を遅いrateで表示、その後もそのrateで表示する	UETのどの時点(ユーザが経験したある瞬間)をtargetとするかを決定する

えない。そこでユーザにとって意味のある rate 変化をあらかじめ設定し、そのパターンに応じて必要となるコントローラを用意したプロトタイプシステムをそれぞれ作成した。表 1は3種類の rate 変化パターンのグラフと、それを利用したタスクシナリオ、特徴、および必要となるコントローラについて示したものである。以下にそれぞれのプロトタイプにおける rate 変化と、利用した空間的なインタラクションについて説明する。

Type1

Type1 は全体を速い速度で見て、流れを掴んだり、全体を遅い速度で見て細かな動きを観察するようなタスクを想定している。rate を常に一定に保つものであるため、rate controller はグラフではなくスライダを用いた。図 7では3種類の TbVP がそれぞれのスライダ型の rate controller により示された値を保持して表示されている。空間的なインタラクションとしては、画面中段部のスペースではユーザが自由にそれぞれの TbVP を配置して見比べることができる(表示位置の変更)。また上段のスペースでは透明度を変化させた TbVP を正確に重ね合わせて置くことができ、空間的な同期を取ることが可能となっている(全領域の同期)。ひとつの TbVP において表示速度が相対的に変化しない表現であるため、卓球の試合やダンスなど単調なリズムを持つ映像からリズムの崩れる瞬間を探すなどのタスクで利用することが考えられる。

Type2

Type2(図 8)はユーザ興味のある部分をゆっくりと観察し、それ以外の部分は流し見をするようなタスクを想定しており、MDT の特定の時点から離れるにつれて速い速度で表示されるものである。MDT の進行を表す Presentation Slider と並行に Target Selector を配置したため、ユーザは着目したい MDT の時点を目安に選択することができる。ひとつの TbVP に表示速度の速い部分と遅い部分を持たせることによって、時間的な前後のつながりを失わずに、特定の時点を詳しく観察することができ、スポーツの一連の試合の流れの中で起こる、際立ったプレーを観察することなどに効果的である。

Type3

Type3(図 9)は、ユーザが TbVP を観察していて気になったところをその直後に再び詳しく見る、といったタスクを想定したものである。ユーザが target とした瞬間に表示されていた映像を見ることが目的であるが、その前後の流れを失わないように、target が設定された直後に高速で一定の時区間を巻き戻すようにしている。その後、再び target 付近に近づくにつれてrateを徐々に遅くし、target を過ぎると徐々に速くして元の rate に戻る。ここでの target は MDT そのものではなく UET においてユーザが興味を持った瞬間であり、ユーザが「今」と思った瞬間に利用しやすいボタンインタフェースのコントローラとした。空間的なインタラクションとしては部分的な同期を行っており、「今」という瞬間だけでなく、どの位置でどの

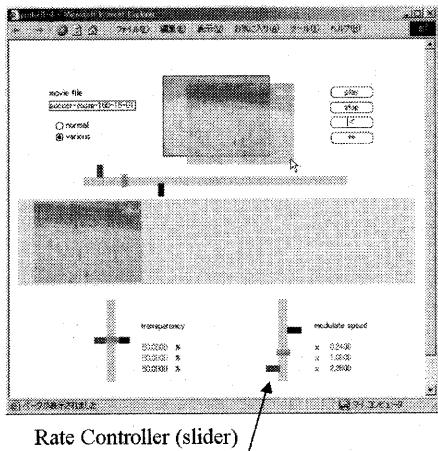


図 7: Type1

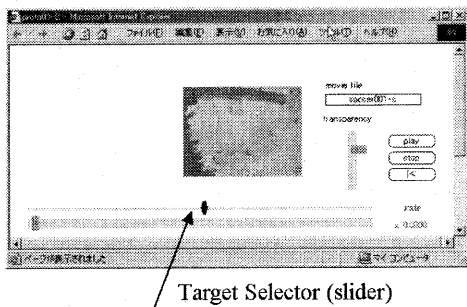


図 8: Type2

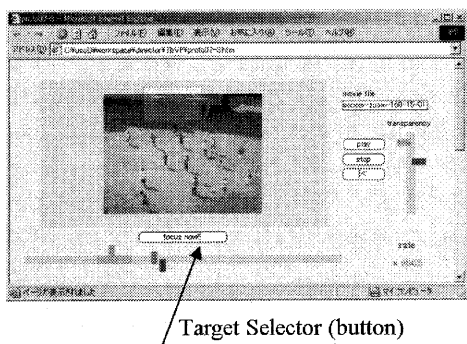


図 9: Type3(1)

大きさの領域の「今」を再び見たいかをインタラクティブに指定できる。加えて時間的な同期も取り入れており、再表示により時間遅れを生じた領域を、周りの映像の時点に追いつかせることが可能であるため、

何度でもインタラクティブに「今」を設定し、見直すことができる。

5 まとめ

本稿では、映像メディアの内容をユーザが理解するために探索的な分析を行うことを想定し、TbVPとして表示される映像に対する空間的、および時間的インタラクションを可能にする環境を構築した。

プロトタイプシステムとして最後に述べた 3 つの TbVP Browser は、ユーザにとって意味のある rate の変化をあらかじめ設定したものであった。今後は、このような rate 変化のパターンをよりインタラクティブに利用することを目指し、これらをギャラリーとして保存しておき、映像を観察している際、必要に応じてユーザが選択的に利用できる環境を構築する予定である。

参考文献

- [1] Nuno Correia, Teresa Chambel, "Active video watching using annotation," Proceedings of the seventh ACM international conference on Multimedia (Part 2), pp.151-154, Orlando, Florida, United States, 1999, ACM Press
- [2] Kumiyo Nakakoji, Akio Takashima, Yasuhiro Yamamoto, "Cognitive Effects of Animated Visualization in Exploratory Visual Data Analysis", Proceedings of Information Visualisation (IV01), pp.77-84, London, United Kingdom, 2001, IEEE Computer Society
- [3] 佐藤隆, 阿久津明人, 南憲一, 外村佳伸, "Coaster: 映像の時空間直感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, 1999, 情報処理学会
- [4] Natalia Andrienko, Gennady Andrienko and Peter Gatalsky, "Supporting Visual Exploration of Object Movement," Proceedings of the Working Conference on Advanced Visual Interfaces(AVI2000), pp.217-220, Palermo, Italy, 2000, ACM Press
- [5] Francis C. Li, Anoop Gupta, Elizabeth Sanocki, Li-wei He and Yong Rui, "Browsing digital video," Proceedings of Human Factors in Computing Systems (CHI2000), pp.169-176, Hague, Netherlands, 2000, ACM Press
- [6] QuickTime <http://www.apple.com/quicktime/>