

個人ビデオのための眼球運動を利用したデジタルカメラワーク

中村 亮太[†] 井上 亮文^{††}
岡田 謙一^{†††} 市村 哲^{††}

近年、デジタルビデオカメラやブロードバンドの普及が進み、様々なシーンでビデオカメラを利用する機会が増え、また Web サイトで撮影した映像を公開することも行われている。配信する映像は視聴効果の高い、魅力あるものであることが望まれるが、素人には難しく、映像制作を簡単に行える技術が期待されている。そこで著者らは眼球運動を用いて、だれにでも容易に視聴効果の高い映像へ編集することが可能なシステムの構築を行った。本システムでは映像視聴時の眼球運動と仮想的なカメラワークを対応させることで、自然なカメラワークが加わった映像へ自動的に編集することが可能である。評価実験の結果、素人が撮影した映像よりも視聴効果の高い映像を作成することができ、構築したシステムの有用性を確認した。

A Digital Camera Work Using Gaze for Amateur's Videos

RYOTA NAKAMURA,[†] AKIFUMI INOUE,^{††} KEN-ICHI OKADA^{†††}
and SATOSHI ICHIMURA^{††}

Recently, the digital video camcorder has become popular, so we have gotten an opportunity to videotape a daily life and we have showed our videos on the home page of a web site or the video-blogging. However it has difficult to edit videos for amateurs. Consequently, they have expectations of automatic video editing system. In this paper, we constructed an automatic video editing system using gaze. We related an eye-movement while we watch videos to camera works such as panning or zooming so that our System can add camera works to amateur's videos effectively. As a result of comparison amateur's videos with our videos, our videos were more attractive than other videos.

1. はじめに

デジタルビデオカメラの普及が進み、子供の運動会、学芸会、旅行やスポーツなど、様々なシーンでビデオカメラを利用する機会が多くなった。また、ブログや Web サイトで撮影した動画を公開し、動画を通してコミュニケーションを広げていくことや^{1),2)}、さらには同じイベントを撮影したユーザ同士が映像を共有し、Web 上で編集を行うことも考えられている³⁾。このように動画を扱うコミュニティが急速に拡大しており、新たな社会システムの 1 つとして今後も成長していくことが予想される。

このような状況において、制作される映像は視聴効果の高い、魅力ある映像であることが望まれるが、素人が撮影すると単調で退屈な映像になりがちである。また、撮影者は撮影作業に集中しなければならないため、自分自身がイベントに参加することや、一観客として楽しむことができない。そこで、自動的に撮影、編集を行う技術が注目されている。従来技術として、自動的にカメラワークや要約を行うシステムが存在し、スポーツや料理、講義映像などの分野で利用されている。しかし、これらのシステムは、撮影対象に特化しているため、別の対象を撮影しようと思えば、新しくシステムを構築しなければならず、汎用性に欠けるとい問題がある。

そこで、本研究では眼球運動を用いることによって、様々なシーンにも利用ができ、だれにでも簡単に扱え、なおかつ魅力的な映像へ自動的に編集することができる手法を考案し、プロトタイプシステムの構築を行った。本システムでは事前に撮影した映像を眼球運動を測定しながら視聴するだけで、自動的にカメラワーク

[†] 慶應義塾大学大学院理工学研究科
Graduate school of Science and Technology, Keio University

^{††} 東京工科大学コンピュータサイエンス学部
School of Computer Science, Tokyo University of Technology

^{†††} 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University

の加わった映像を作成することができる。そのため、撮影時には自らカメラワークを行って撮影する必要なく、また、高度な編集技術を持たない素人にも本システムを簡単に扱うことができる。評価実験の結果、本システムを利用した映像は、アマチュアカメラマンが撮影した映像よりも、パンやチルト、ズームなどのカメラワークが適切に加わっており、視聴効果の高い映像を作成することができた。

2. 背景

2.1 ホームビデオ制作における問題点

ビデオカメラのデジタル化にともなう高画質化・小型化・軽量化に加え、パソコンに取り込んだ編集・ホームページやブログへの活用など用途が広がったことで、ビデオ撮影をする機会が増えてきた。しかし、ビデオカメラの性能が向上しても、優れたカメラワークの加わった質の高い映像を撮影するには、高度な撮影技術や経験を身につける必要があり素人には難しいといえる。さらに撮影作業自体が撮影者の行動を制限してしまう。多くの撮影者はビデオカメラを持ち、より良い映像を撮影するためにカメラを操作するが、撮影作業が長時間にわたる場合、負担が大きく、また、カメラを操作し続けなければならないため、パーティーやスポーツの試合など自分自身がイベントに参加することや、一観客として楽しむことができないという問題があげられる。そこで、自動的に撮影・編集を行う技術が注目されている。

2.2 従来技術の問題点

従来からデジタル技術を導入してプロのカメラマンや編集者を必要としない自動映像生成についての研究が注目されている。従来の自動撮影システムの多くは首振りカメラや複数のカメラをスイッチングさせて撮影するシステムである^{4)~6)}。一方、画角の広い固定カメラで撮影した映像から表示領域を限定し、この領域を動かすことで仮想的にカメラワークを加える研究もある^{7)~9)}。ここで行われているように講師やサッカーボールを追尾するだけの映像を作成するのであれば、従来システムのように動体を検出することで自動的に撮影することが可能である。しかしそれ以上の映像内容の理解を画像認識などで自動的に行うことは非常に難しく、他の手法が必要である。一方、著者らのシステムでは、計算機による画像認識を行わず、鑑賞者の視線を利用するアプローチを行っている。鑑賞者の視線方向をカメラの撮影方向と一致させることで、鑑賞者の目に止まったあらゆる対象を映し出すことが可能であり、従来システムとは異なる特徴を有してい

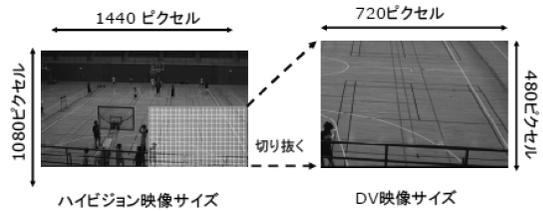


図 1 仮想カメラワーク

Fig. 1 A virtual camera-work.

る。たとえば、サッカーの試合などではボールをとらえるだけでなく、ボールを保持していない周りの選手をとらえることもできる。また、鑑賞者がある撮影対象を注視するようなシーンがあったとき、注視部分をズームすることであたかもカメラマンがカメラワークを行ったかのような自然な映像を作り出すことができる。以上のように本システムでは映像内容を理解したうえで映像を作成することができるため、様々な種類のビデオに適用可能であると考えられる。

2.3 デジタルカメラワーク

デジタルカメラワークとは実際に撮影時にカメラマンがパンやチルトなどのカメラワークを行うのではなく、撮影後の映像の見せ方を工夫することであたかもカメラワークが加わったかのような映像を作り出す技術である^{7),8)}。具体的には、HDV (High-Definition Video) カメラで撮影され、 $1,440 \times 1,080$ ピクセルのハイビジョン映像サイズで保存された映像の中から一部分を切り出し、移動させることで擬似的にカメラワークのある映像を作成する(図1参照)。切り出した領域を左右に移動させることでパンニング、上下に移動させることでチルティング、表示領域の拡大率を変化させることでズームの効果を作り出すことができる。また、切り出し位置やフレームサイズを瞬間的に変化させることで仮想的なカメラのスイッチングを実現することができる。このようなデジタルカメラワークはデジタルシューティングとも呼ばれており⁹⁾、講義やスポーツ映像に適用されている。本論文でもこの技術を利用し、高解像度なハイビジョン映像から映像を切り出すことで、カメラワークが加わった DV (Digital Video) サイズ (720×480 ピクセル) の映像を作成することにした。

3. 眼球運動

視覚刺激を見ているときの眼球運動が心的な処理とどの程度対応しているかは未解明な部分が存在するが、眼球は機械的な運動やランダムな運動を行っているのではなく、何らかの心的処理と連動して反応するもの

と考えられている¹⁰⁾。たとえば、視覚刺激の中で情報価値が高い部分に対して比較的長い注視が行われることが知られている¹¹⁾。これらのことから眼球運動をカメラワークに利用すれば人間の心的処理をカメラワークに反映させることが可能であると思われる。

3.1 眼球運動の特徴

眼球運動は注視対象の移動の仕方により異なっており、固視微動、随従運動、跳躍運動の3つに分類される¹²⁾。固視微動は注視時に生じ、非常に小さな動きであるため視覚系にとってノイズと見なされることもあるが、網膜状の解像度を保ち、はっきり見るために欠かせない運動と考えられている。人は一点を注視しているとき、目の動きは止まっているように見えるが、実際には小さな動きをたえず行っている。眼球運動データから視覚情報を受容している注視の状態を抽出するためには、一定の定義が必要である。広く用いられている定義は眼球運動の速度が一定以下 ($5 [\text{deg}/\text{sec}]$) に、もしくは視線の一定範囲内 ($2 [\text{deg}]$ 以内) にとどまっている状態が一定時間以上 ($150 [\text{ms}]$) 持続した場合とするものである。

随従運動は動いている対象を追従しているときに生じる滑らかな眼の動きのことである。SPM (Smooth Pursuit Movement) とも呼ばれ、最高速度が $30 [\text{deg}/\text{sec}]$ の滑らかに動きであり、主に動いている視標を追従しているときに生じる。

跳躍運動はサックードとも呼ばれ、跳ぶような速い動き、本を読んでいるときや、日常様々な視覚対象を次々と注視するときに生じ、一般的に視点の移動時に見られる。視標の速度が遅いときは、随従運動で視標を遅れなく追跡している。しかし、視標の速度が速くなると、もはや随従運動で追えなくなり、視標に対して遅れていく。跳躍運動はそのようなときにも生じ、遅れた分を取り戻すべく、跳ぶような速い動きを行う。跳躍運動の速度は $600 [\text{deg}/\text{sec}]$ にも及ぶといわれている。

3.2 眼球運動の測定方法

本論文では角膜反射法を利用して眼球測定を行った。角膜反射法は角膜上の赤外 LED の虚像が、角膜と眼球の回転中心の違いにより、眼球運動にともなって平行移動するのを、赤外感度を持つビデオカメラで検出するものである。角膜上の虚像はビデオカメラで確認でき、センサの装着が容易である。また、時間的・空間的解像度はカメラの解像度とフレーム周波数に依存する。

4. 視線によるカメラワーク

著者らは映像視聴時の眼球運動を測定し、視線の軌跡を基にデジタルカメラワークを行うことを試みた。



図2 眼球測定実験風景

Fig. 2 An image of experiment for measuring eye-movement.

眼球運動データからカメラワークを行うために、まず映像視聴時の視線の軌跡と映像との対応を分析した。

4.1 眼球運動の測定方法

従来研究において特に実験対象とされることの多かった球技スポーツ、会議の2種類の映像を用意した。つねにボールの動きがともなう球技スポーツの例として大学サッカーサークルの練習風景 (video1)、身体の動きはあまりなく多人数が参加する会議のような風景として研究室の懇談会の様子 (video2) を撮影し、大学生の被験者3名に眼球測定器を装着させながら30分間視聴させた。視聴させた映像はHDVカメラによって撮影した俯瞰映像であり、カメラワークのない固定映像である。被験者には画面から $1 [\text{m}]$ 離れた位置から映像を視聴させた。

眼球測定器にはナックイメージテクノロジーのアイマークレコーダ EMR-8B を用いた (図2 参照)。EMR-8B では、帽子に視野カメラとユーザの眼球を撮影する赤外線カメラが備わっており、付属された解析ソフトにより、リアルタイムにデータが解析される。データは $60 [\text{Hz}]$ でサンプリングされ、瞳孔の径の大きさ $[\text{mm}]$ 、視線の座標 $x [\text{deg}]$ と $y [\text{deg}]$ の3つデータが出力される。本実験では眼球測定が極力映像視聴の妨げにならないように片眼で測定を行った。

4.2 実験結果

図3、図4は映像視聴時の眼球運動の例として被験者Aの視線座標 $(x, y) [\text{deg}]$ と瞳孔の径 $[\text{mm}]$ の変位が示されている。グラフは視野映像の中心を0として表されており、グラフが右下がりとなれば、画面の左側に視線が移動したことを表している。なお video1、2ともに撮影対象の動きが主に水平方向であったため、図のように特に x 座標に大きな変化が表れている。

video1 視聴時の視線座標は、ボールを目で追う随従

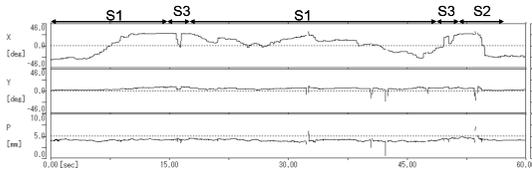


図3 video1を視聴しているときの眼球運動(例)
Fig. 3 A case of eye movement of watching video1.

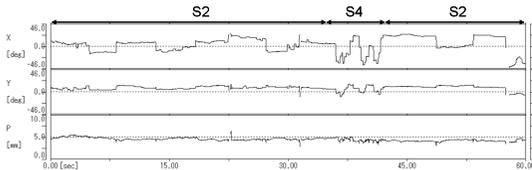


図4 video2を視聴しているときの眼球運動(例)
Fig. 4 A case of eye movement of watching video2.

運動が中心であり図3中S1部分のように視線の座標が徐々に変化している。このように指標の速度が遅いときは、随従運動で遅れることなく追跡できるが、移動距離が長く素早いパスやシュートが起きた場合、随従運動で追えなくなり、視標に対して遅れていった。この遅れた分を取り戻すべく、跳ぶような速い動きをする跳躍運動が行われていた。このとき、視線は大きく移動したため、図3中S2部分のように大きく下降した。また、図3中S3部分ではボール追従時に、瞬時に視線を他へ移動し、再びボールを追従するという視線の動きが観察された。これは周りの選手の動きやゴールなどの位置関係を把握するために瞬時に視線を移動させたのではないかと考えられる。また、ボールが飛んでいく方向を予想したときにも無意識的に視線を移動することがあり、興味深い特徴であった。被験者間で視聴の違いが表れるか観察したところ、ボールを追従することが多い者、ボールがない部分へ頻りに視線を移動する者が存在することが分かった。

video2の映像は、複数の箇所では会話がそれぞれ独立して行われており、video1のサッカーボールとは異なり、注目しやすい対象が複数存在していた。また懇談会の参加者は椅子に着席していたため、ボールのような連続的な指標の移動は行われていなかった。このようにvideo2では、随従運動よりも固視微動と跳躍運動を繰り返す様子が多く観察された(図4中のS2部分)。これはvideo1で素早いボールの動きを目で追ったときに表れた様子と類似する。具体的には、ある人物に注視し始めると、視線の座標変化はほとんどなくなるため、グラフのように水平な直線を描く(固視微動)、次に別の人物へ視線を急激に変化させたため、グラフは短時間に大きく上昇、下降する(跳躍運動)。video2

において、この一連の動きを繰り返したときの視対象は、現在会話を行っている人物であったり、離れた場所に位置する人物であったりしたことが観察された。また図4中S4のように跳躍運動が連続的に行われることが観察された。これは画面全体を素早く見渡し、次の視対象を探すという作業の現れであると考えられる。

4.3 眼球運動とカメラワーク

<パン・チルト>

本システムで実装した視線データの使用方法を述べる。撮影した映像から画像切り出し領域(DVサイズ: 720×480ピクセル)を上下左右に移動させることで仮想的なパンとチルトを実現する。前節の図3中S1のような視線の移動があった場合、これに合わせて映像の切り出し領域を移動させる。このとき、切り出し領域の中心座標が視線座標と一致するように設定した。しかし、視線の座標の変化は不安定であるため、つねに一致させてしまうと、映像のブレが目立ち非常に見にくい映像となってしまふ。そこで後述するような閾値を設け、一定量の座標移動がなければ切り出し領域を変化させないようにした。

上記のように撮影対象に注目した視線がどの程度移動したとき、それを追尾するようにカメラワークを行うべきかを調べるために、まず、カメラワークを水平・垂直方向で分け、それぞれに対して異なるパラメータで設定した映像(5分間)を複数用意し、被験者5名に視聴させ、各アンケート項目に対して5段階で評価を行わせた。

図5、図6は視線の水平・垂直方向の移動量がある閾値を超えたときにパン・チルトを行う設定で動画を作成したものを被験者によって評価させた結果である。 x 軸はそれぞれの方向について閾値[ピクセル]を変化させたものであり、 y 軸は被験者による評価の平均点である。アンケートの結果、水平方向では200[ピクセル]、垂直方向では150[ピクセル]と設定したとき、評価項目(1)の点数は最高得点を得た(水平方向: 平均点 = 3.5 標準偏差 $\sigma = 0.52$, 垂直方向: 平均点 = 3.4, $\sigma = 0.51$)。一方、評価項目(2)では、水平・垂直方向ともに閾値が小さいときには点数は高いが、閾値を増加させると減少していった。これは閾値を大きくすると、つまり、撮影対象を追う視線の動きが大きくなければ、カメラワークを行わない設定にすると、画面から撮影対象が頻りに外れてしまったことが原因であると考えられる。よって移動量は小さいとき、つまり撮影対象をつねに追尾するようなカメラワークが望ましいように思われるが、そのように設定してしまうと逆に震えるような映像になってしまう。

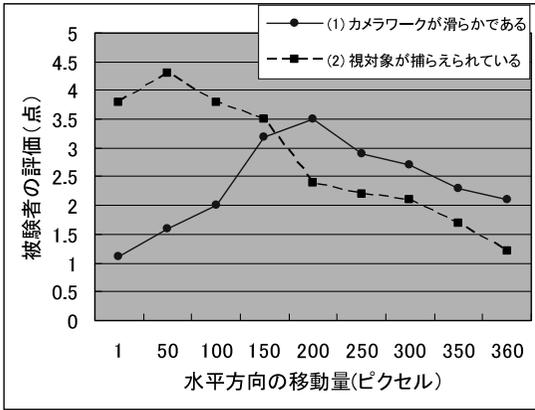


図5 カメラワークの評価結果（水平方向）

Fig. 5 The questionnaire result about camera-work.

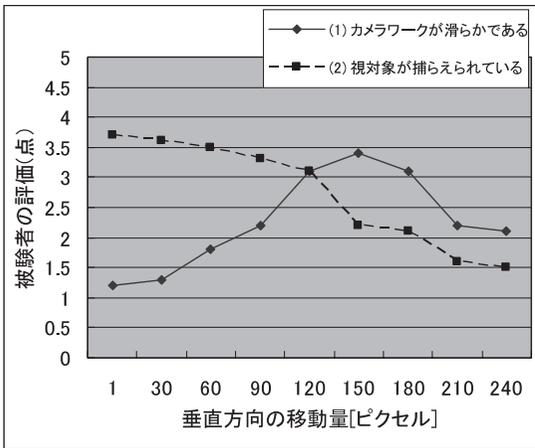


図6 カメラワークの評価結果（垂直方向）

Fig. 6 The questionnaire result about tilting.

評価項目 (1) : パンやチルトは滑らかであったか
 < 5 段階評価 >
 とても滑らかであった : 5, かなり滑らかであった : 4,
 滑らかであった : 3, やや滑らかではなかった : 2,
 滑らかではなかった : 1
 評価項目 (2) : 撮影対象をとらえられていたか
 < 5 段階評価 >
 つねに中心でとらえられていた : 5,
 ほぼ中心で画面から外れることなくとらえられていた : 4,
 画面から外れることなくとらえられていた : 3
 画面から外れることが何度あった : 2
 画面から外れることが多かった : 1

そこでカメラワークが滑らかでかつ撮影対象が画面から外れることがなければよいと判断し、この条件をクリアできたと判断できる評価点 3 点を越えたときを最適なパラメータとして設定することにした。したがって水平方向では、評価項目 (1), (2) に対するそれぞれの平均点が 3.2 ($\sigma = 0.67$), 3.5 ($\sigma = 0.52$) となった 150 [ピクセル], 垂直方向では、(1) 3.1 ($\sigma = 0.32$),

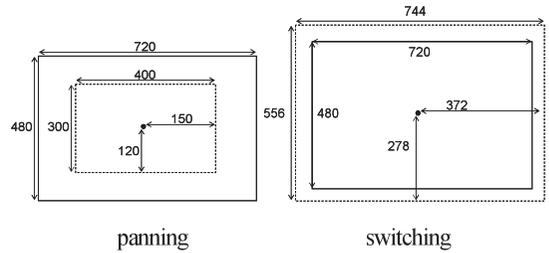


図7 カメラワークの条件

Fig. 7 The area of camera-work condition.

(2) 3.1 ($\sigma = 0.56$) となった 120 [ピクセル] を閾値として設定することにした。

< スイッチング >

4.2 節で述べた S2 のように急激な視線の移動があった場合に対してもパンやチルトを適用させてしまうと速度の速いカメラワークになってしまい非常に見にくい。このような場合、映像をスイッチングすることにした。また S3 のように瞬間的に視線移動の往復があった場合、一瞬映像を切り替えて、瞬時に元の映像に戻すスイッチングを行うことにした。具体的にはサッカー映像を例にした場合、ドリブルをしている選手を表示しているときに、パスを受け取りそうな味方の選手を一瞬表示させるといったカメラワークを行う。

以上のカメラワークを行うために 4.2 節で分析したデータの中で S2, S3 が表れたときの x, y 座標の最小移動量を求めたところ x 座標では y 座標ではこれらの値を超えた場合、切り出し領域の座標を移動の始点から終点まで跳躍させることで仮想的なスイッチングを実現させることにした (図 7 参照)。

< ズームイン・アウト >

図 4 中の S2 部分内で観察されたように一定時間、大きな視線の移動が起こらず、停滞している場合があった。このような固視微動が起きているときは、興味や関心を持って指標を注視している状態であると判断できるため、その指標を強調させるためにズームを行うこととした。具体的には画像切り出し領域を DV サイズ (720 × 480 ピクセル) から徐々に狭めていき、画質の劣化を考慮した DV の 1/4 のサイズ (360 × 240 ピクセル) まで領域を狭めることで仮想的なズームングを実現した。ズームングを行う条件としては、視線の座標が一定時間、限られた領域内にとどまっていた場合、その領域を拡大する。また、ズーム後に大きな視線の移動があった場合、自動的にズームを解除し、少しずつ標準のサイズへ戻るように設計した。

またズームの発生する頻度に関して考察すると、ズーム回数が少なすぎると、映像が単調になってしまい、

逆に頻繁にズームを行うと見にくい映像になってしまう恐れがある。過去の事例では注視判定基準の例¹²⁾として 2 [deg] にとどまっている状態が 150 [ms] 持続した場合、これを注視状態と判定する方法が述べられている。そこで著者らはこのパラメータでズームを行った映像を被験者に視聴させ、評価結果からパラメータの調整を行うことにした。

映像評価では注視条件を変えて作成した映像を視聴させ、「見やすい映像である」という観点から被験者 10 名に 5 段階（あてはまる：5，ややあてはまる：4，どちらともいえない：3，ややあてはまらない：2，あてはまらない：1）で評価させた。停留範囲の値は上記のパラメータをピクセルに換算し（約 25.1 [ピクセル]），その値を中心に变化させたものであり，視線の移動量がそれぞれ 10，15，20，25，30，35，40 [ピクセル] 未満であった場合で比較した。

調査の結果，閾値が 25 [ピクセル] 未満の場合，評価が最も高く（平均 4.2，標準偏差 0.42），作成された映像ではズームが効率良く加えられており，見やすく面白みのある映像であったという意見を多く得た。しかし 20 [ピクセル] 以下になるとズームがほとんど行われず単調さが目立つようになり，また 30 [ピクセル] 以上の場合には，ズームの発生が過剰に起きてしまい，映像酔いを起こすような不快な映像であるという意見が多く得られた。よって，本手法では，20 以上 25 [ピクセル] 未満でズームを行うことを条件とした。

また，ズームを行うスピードについても検討を行った。映像のフレームサイズが 720×480 から 360×240 まで変化するのに要する時間をそれぞれ 1，2，3，4，5 秒間で設定した映像を作成し，被験者に 5 段階の主観評価を行わせた。その結果，2 秒間（1 フレームあたり x 軸：±3 [ピクセル]，y 軸：±2 [ピクセル]）で变化させたときの映像の評価が最も高いことが分かり（平均 3.8，標準偏差 0.33），この値をズームイン・アウトの適正スピードとした。

<全体表示>

視聴者は見たい対象物を探す作業を行うとき，4.2 節の S4 のように跳躍運動を繰り返し，画面全体を連続的に大きく走査するように視聴する。このような眼球運動が起きたときは画面切り出し領域を一時的に HDV サイズ（1,440×1,080）まで拡大し，S4 の終了と同時に DV サイズに縮小することにした。4.2 節の S4 発生時の視線の総移動量について調査したところ，平均で 1,125.8 [deg/sec] であったため，この値を全体表示の条件とした。

5. システム評価

5.1 評価実験方法

本システムによって映像に加えられたカメラワークが適切であるか，または，様々なビデオシーンで本システムを使用することが可能であるかを検証するために，大学サッカーサークルの練習風景と研究室室内での懇談会の様子を撮影対象とし，アマチュアカメラマン（大学卒研究生：男 4 名，女 1 名）が撮影した映像と同カメラマンが本システムを利用して作成した映像を比較した。撮影方法はアマチュアカメラマンに DV カメラを持たせ，自由に撮影を行わせた。また，本システムを利用して映像を作成するために，三脚で固定した HDV カメラで俯瞰映像を撮影しておいた。なお，アングルの違いによって映像評価に影響を及ぼす恐れがあったため，カメラマンと HDV とは同位置から撮影を行わせた（計 20 サンプル×10 分間）。

本システムによる映像の作成作業では，撮影した俯瞰映像を大型ディスプレイに表示し，眼球運動を装着した被験者に視聴させ，測定した眼球運動データを基にカメラワークの加わった映像を作成した。映像の評価作業では，被験者 10 名（大学卒研究生：男 9 名，女 1 名）を 2 つのグループに分け，1 つのグループにはアマチュアカメラマンが撮影した映像を先に視聴させ，その後で本システムが作成した映像を視聴させた。また，もう一方のグループにはこれとは逆の順序で映像を視聴させた。そして，表 1，表 2 に示すように 5 つの評価項目について 5 段階の主観評価を行わせた。これをサッカーの映像と懇談会の映像の 2 種類で繰り返した。表 1，2 に評価結果を示す。

5.2 実験結果と考察

表 1，2 中の Video-I 群はアマチュアカメラマンによる撮影映像であり，Video-II 群は本システムによって作成した映像である。表中には 5 つの評価項目に対する被験者間の平均値（ $n = 10$ ）が示されており，また，Video-I 群，Video-II 群の 2 つ平均値に有意差があるかどうか調べるためにノンパラメトリック検定である Wilcoxon 順位和検定¹³⁾を行った。評価実験の結果，2 種類のコンテンツともにすべての評価項目で Video-II 群の平均値は Video-I 群の平均値を上回る結果を得ることができた。また，Wilcoxon 順位和検定の結果，2 種類のコンテンツともに全評価項目において有意水準 5% 以下で平均値に差があることを確認した。特に評価項目 ①，② のパン・チルトやズームのよし悪しに差が出た。Video-I 群の映像はパンやチルトの動きが不安定であり，またアマチュアカメラ

表 1 サッカー映像の評価実験
Table 1 The evaluation result of soccer videos.

	Video-I 郡	Video-II 郡	Wilcoxon
	平均値	平均値	順位和検定 p 値
①	2.76	3.72	0.000413***
②	2.24	3.66	0.000151***
③	2.34	3.40	0.003216**
④	2.88	3.38	0.024548*
⑤	2.62	3.20	0.012572*

p<0.001***, p<0.01**, p<0.05* (n = 10)

表 2 懇談会映像の評価実験
Table 2 The evaluation result of talkfest.

	Video-I 郡	Video-II 郡	Wilcoxon
	平均値	平均値	順位和検定 p 値
①	2.36	3.56	0.002404**
②	2.14	3.84	0.000671***
③	2.90	3.60	0.029948*
④	2.56	3.20	0.016509*
⑤	2.34	3.16	0.010102*

p<0.001***, p<0.01**, p<0.05* (n = 10)

- ① パン・チルトが滑らかで見やすいカメラワークである
② ズームイン・アウトのスピードやズーム率が的確である
③ 何に注目すればよいか分かりやすい
④ 見ていて飽きない、また、魅力的な映像である
⑤ また視聴したいと思う

< 5 段階評価 >

あてはまる : 5, ややあてはまる : 4,

どちらともいえない : 3,

ややあてはまらない : 2, あてはまらない : 1

マンは非常に素早くパンやチルトを行う傾向があり、見にくい映像であったのに対し、Video-II 群の映像は激しいカメラワークは発生せず安定していた。以上のように、本システムで作成した映像の方がより滑らかで見やすいカメラワークを加えることができることを確認した。

次に評価項目 ③「何に注目すればよいか分かりやすい」に対して、被験者から分かりやすい(平均点 3.5)という回答を得ることができ、この結果から、作成者の意図したカメラワークをある程度精度良く再現することが可能であることを確認した。また、評価項目 ④, ⑤ の結果に示したように Video-I 群より面白みがあり、魅力的であることも被験者の意見から確認できた。

映像の種類の違いによる評価に差が出たか考察した

ところ、2つのコンテンツともに Video-I 群の映像よりも適切なカメラワークを加えることができた。懇談会映像のようにボールなどの客観的指標が1つでない場合でも、視線の移動と意図したカメラワークがまったくの1対1で対応することはないものの、加えられたカメラワークには鑑賞者の興味や関心が反映されていた。したがって、映像中にはっきりした1つの注目対象が存在しない場合でも、本提案手法が適用可能であると考えられる。

しかしながら次のような改善点も明らかになった。まず、視線データからカメラワークを滑らかに行えるようにいくつかの処理を施したが、それでも若干映像のブレが発生してしまうことがあった。これは眼球運動において随従運動から跳躍運動に切り替わるときに起こる誤差であると考えられる。速度の速い指標に対し、随従運動で追えないとき、跳躍運動で遅れを取り戻すように素早く視線を移動する試みが行われる。このとき眼球では指標を正確にとらえ続けることはできず、結果、映像ブレを引き起こす原因になった可能性が高い。今後、さらに視線データの分析を進め、様々な状況に応じた視線データの処理を行う予定である。そのほか、眼球測定にあたってはユーザの自由度が大きい非接触型の測定器を利用することが望ましく、より実用性のあるシステムを構築する場合にはこのような測定器を利用することが理想的である¹⁴⁾。

6. おわりに

本論文では、新たに社会システムの1つとして拡大していくことが予想される動画のコミュニティを支援するシステムの開発を行った。本システムによって素人でも映像視聴時の眼球運動を利用することで自然なカメラワークを仮想的に加えることができ、アマチュアカメラマンが撮影した映像よりも視聴効果の高い映像を作成することができた。本システムのような眼球運動から映像を作成する技術は、個人ためのビデオ制作だけでなく、映像を通して学習を行う教育現場にも応用が可能であり、様々な分野で利用可能であると考えられる。

参考文献

- 1) FlipClip. <http://www.flipclip.net/> (2006/5/1 アクセス)
- 2) Broadband Portal Platform. <http://www.ntt.com.co.jp/bpp/> (2006/5/1 アクセス)
- 3) Ichimura, S. and Matsushita, Y.: Web-based Video Editing System for Sharing Clips Collected from Multi-users, *The 7th*

IEEE International Symposium of Multimedia, pp.104-111 (Dec. 2005).

- 4) 錦織, 菅沼, 谷口: 黒板講義を対象とした遠隔講義のための講義自動撮影システムの構築, 情報処理学会 Vol.2000, No.95, pp.63-80, 2000-CE-57 (2000).
- 5) Kameda, Y., Ishizuka, K. and Minoh, M.: A Live Video Imaging Method for Capturing Presentation Information In Distance Learning, *Proc. ICMCS'99*, Vol.2, pp.897-902 (1999).
- 6) 大西, 影林, 福永: 視覚情報統合による会議映像の自動撮影, 電子情報通信学会論文誌, Vol.J85-D-II, No.3, pp.537-542 (2002).
- 7) 大西, 泉, 福永: デジタルカメラワークを用いた自動映像生成, 画像の認識・理解シンポジウム, pp.I-331-I-336 (July 2000).
- 8) 横井, 藤吉: 高解像度映像からの自動講義ビデオ生成—仮想カメラワークの実現, 第11回画像センシングシンポジウム (June 2005).
- 9) 熊野, 岩本, 有木, 塚田: ボールと選手に着目したデジタルカメラワークの実現法, 画像の認識・理解シンポジウム (MIRU2004) (July 2004).
- 10) 中島義明: 映像の心理学—マルチメディアの基礎, pp.167-168, サイエンス社 (2004).
- 11) 三浦利章: 視覚的行動・研究ノート—注視時間と有効視野を中心として, 大阪大学人間科学部紀要, 8, pp.171-206 (1982).
- 12) 福田忠彦: 生体情報システム論, 産業図書 (1995).
- 13) 菅 民朗, 檜山みぎわ: 初めて学ぶ統計学, 現代数学社 (2003).
- 14) 大野健彦, 武川直樹, 吉川 厚: 2点補正による簡易キャリブレーションを実現した視線測定システム, 情報処理学会論文誌, Vol.44, No.4, pp.1136-1149 (2003).

(平成 18 年 5 月 30 日受付)

(平成 18 年 11 月 2 日採録)



中村 亮太 (学生会員)

慶應義塾大学大学院理工学研究科開放環境科学専攻。2002年3月東京工科大学工学部卒業, 2004年3月同大学大学院前期博士課程修了。現在, 慶應義塾大学大学院後期博士課程在学中および東京工科大学 Linux オープンソースセンター嘱託研究員。生体情報工学, マルチメディア処理の研究に従事。DICOMO2004 優秀プレゼンテーション賞受賞, DICOMO2005&2006 最優秀プレゼンテーション賞および優秀論文賞受賞, 平成 18 年度山下記念研究賞受賞。



井上 亮文 (正会員)

1999年慶應義塾大学理工学部計測工学科卒業。2001年同大学院前期博士課程修了。2005年同大学大学院後期博士課程修了。博士(工学)。現在, 東京工科大学コンピュータサイエンス学部助手。マルチメディアオーサリング, 実世界志向インタフェース, ネットワークセキュリティの研究に従事。ACM, ヒューマンインタフェース学会各会員。



岡田 謙一 (フェロー)

慶應義塾大学理工学部情報工学科教授, 工学博士。専門は CSCW, グループウェア, ヒューマン・コンピュータ・インタラクション。『ヒューマンコンピュータインタラクション』(オーム社), 『コラボレーションとコミュニケーション』(共立出版)をはじめ著書多数。情報処理学会誌編集主査, 論文誌編集主査, GW 研究会主査等を歴任。現在, 情報処理学会 GN 研究会運営委員, BCC 研究グループ幹事, 日本 VR 学会 CS 研究会副委員長。1996年および2001年情報処理学会論文賞, 2000年情報処理学会 40周年記念論文賞, 日本バーチャルリアリティ学会サイバースペース研究賞, IEEE SAINT '04 最優秀論文賞受賞。電子情報通信学会, 情報処理学会フェロー, IEEE, ACM, 人工知能学会各会員。



市村 哲 (正会員)

1989年慶應義塾大学理工学部計測工学科卒業。1994年同大学大学院理工学研究科博士後期課程修了。博士(工学)。同年富士ゼロックス(株)入社。1997~1999年富士ゼロックスパロアルト研究所(FXPAL)駐在。2002年より東京工科大学助教授。グループウェア, ネットワークサービス, 生体情報活用等の研究に従事。『IT TEXT 基礎 Web 技術』, 『IT TEXT 応用 Web 技術』(オーム社)。DICOMO2003, 2005, 2006 優秀論文賞受賞。ACM, 電子情報通信学会各会員。