

# 画像センシングによる視聴映像の注目度推定システム

片瀨 典史<sup>†</sup> 山本 哲也<sup>†</sup> 藪内 勉<sup>†</sup> 中山 丈二<sup>†</sup> 下倉 健一朗<sup>†</sup>

**あらまし** 映像コンテンツの定量的評価を目指し、画像センシングで得られた視線や瞬目などの情報から、視聴映像コンテンツに対する注目度合いを推定するプロトタイプシステムを試作した。本システムは、カメラや各種センサを用いて映像コンテンツ視聴時における人の反応をセンシングし、その収集時刻と結びつけて収集手段ごとに構造化した当該情報を非同期で蓄積する情報蓄積サーバ、および情報蓄積サーバに蓄積された情報を取り出し、映像コンテンツと当該情報の時間同期をとって再生・表示する機能を持つ同期再生サーバで構成される。情報蓄積サーバと同期再生サーバの連携によって、映像コンテンツに対する視聴者反応の事後同期解析が可能となる。本論文では、映像およびセンシング情報の時間同期表示機能の有用性について報告する。

## Prototype of an Attention Level Estimation System for Video Content Using Image Sensing

Norifumi KATAFUCHI<sup>†</sup>, Tetsuya YAMAMOTO<sup>†</sup>, Tsutomu YABUUCHI<sup>†</sup>, Jyouji NAKAYAMA<sup>†</sup>, and Ken-ichiro SHIMOKURA<sup>†</sup>

**Abstract** In order to evaluate the quality of a video content quantitatively, the authors have developed a prototype system which estimated the viewer's attention level to the video content from gaze and eye blink information by image sensing. The system is composed of an information accumulation server that stores the combination of the sensing information and the collection time using the XML-based format, and a time-synchronous display server. The post synchronous analysis of the viewer's reaction to the content is realized by the cooperation of both the above servers. Experimental results show the usefulness of the time-synchronous display function between the content and the sensing information in this system.

### 1. まえがき

無線 IC タグに代表される情報処理デバイスの小型化・低廉化が急速に進む中、家庭、店舗、オフィスなどの屋内空間に加え、駅や街頭などのオープンなパブリック空間に至るまで、カメラや温度計といったセンシングデバイスが普及しつつある。例えばカメラは、防犯用途に、街頭では遠隔監視や防災などの用途に利用されている。種々のセンサを用いて周囲の環境を認識した結果は、そこにいる人の行動決定や判断に対して影響を与える可能性が高く、重要な情報となりうる。

周囲の環境を、環境に存在するセンサ群を用い

て認識することで人の行動を支援しようとする研究はこれまでも活発に行われてきた。例えば、生活空間である部屋で人の情報を蓄積し人の行動支援に利用する Robotic Room [1] や、複数の知覚エージェントからなる全方位カメラ [2] の研究などがある。一方で人、コンピュータ、ネットワークに適応したサービス提供を目指し、サービスやネットワークを構成する「機能」を最小単位とし、「機能」を動的に組み合わせることでサービスを動的に合成し提供するインターネットアプリケーションプラットフォーム STONE の研究 [3] なども行われている。

これらの研究はセンサを用いた環境認識、人の

<sup>†</sup>日本電信電話株式会社 サイバーソリューション研究所

<sup>†</sup>Cyber Solutions Laboratories

Nippon Telegraph and Telephone Corporation

行動情報収集,あるいはユーザに対する適応的なサービス合成・提供という点で先行しているが,収集あるいは認識した結果としての環境情報や人の情報を一旦蓄積し,そうした履歴情報を加工,再利用するという観点では十分な検討がなされていない。

これらの課題を解決すべく検討を重ねた結果,筆者らは以下に述べる考えを持つに至った。

- 人に関わる過去のセンシング結果や行動履歴情報は,人の行動を詳細に観察・分析したり,心理的状态を推定したりする上で有効である
- 環境や人などの情報の多様性を扱うためには,センシングで対象から取得される種々の情報を自由に組み合わせ可能とする必要がある

したがって本論文では,人の日常生活行動の一例として映像コンテンツ視聴に焦点をあて,映像視聴時における行動や心的状態の情報を収集・蓄積・活用する視聴質調査システムを提案する。この「視聴質」とは「映像コンテンツがどのように見られているか」という情報である。類似研究として,清水らはリモコン操作情報と顔の向き情報という2つのマルチモーダル情報を用いて,「視聴者がどの程度関心をもってテレビを見ているか」という関心度を推定するシステム [4] を提案している。この研究はマルチモーダル情報統合というアプローチに近いが,筆者らの定義する注目度(詳細は第3章で後述)は興味・関心の度合いを扱わないことから,そもそも目的や到達目標が異なる。

本研究では,上述の「視聴質」をコンテンツ制作事業者にフィードバックし,コンテンツ制作の質的な向上を図ることを目的として,後述するフィールド情報流通プラットフォーム FDC (Field Data Center) を利用して,映像視聴時における人の行動情報を収集・蓄積・活用することをねらう。この FDC は,

- 機械やセンサ,人の情報を統一的に扱える
- 一般的なセンサネットワークは単一または限定された種類のセンサを使っているのに対して,FDML を利用すれば複数の物理的なセンサを組み合わせることで論理センサを構成できる
- 履歴を蓄積しておき,蓄積された情報とアプリケーションを柔軟に組み合わせることにより,情報に付加価値をもたせて加工・再利用できる

を大きな特長としている。

筆者らは,情報の取得対象を従来の製造機械ではなく人を中心にすることで,見守りや介護,生活支援等の別の領域でも FDC を十分に応用展開できると考えており,その一環として FDC をベースとした視聴質調査システムの検討を行った。本稿では,その結果や得られた知見について報告する。

## 2. フィールド情報流通プラットフォームを利用した注目度推定システム

### 2.1. フィールド情報流通プラットフォーム

製造機器から取得される多種多様な情報を統一的に扱い構造化するために,筆者らは文書構造を比較的簡易なフォーマットで記述でき,独自のタグを定義可能な XML (eXtensible Markup Language) をベースとした製造業向けデータ記述言語 FDML [5] を提案し,FDML を用いた製造業向け情報流通プラットフォーム [6] を開発してきた。製造業分野では,機械を対象として様々な挙動をモニタしたり推定したりする用途で利用されている実績がある。

#### 2.1.1. FDML

各種機器では,リアルタイムにフィールドデータが生成・出力されるので,データが発生した時刻が重要になる。さらに複数の機器や複数種類のデータを扱う場合には,得られたデータが同じ現象に起因しているか否かというデータ間の関係が重要になる。データ間の時間的関係性を利用して,機器の障害原因の推定やトラブルシューティングに利用するために,FDML では得られた全データに対してミリ秒単位の時刻情報を付加し管理する。

さらに,前もって規定することが困難なデータの名称やデータ型,定義域といったプロパティを FDML 内に閉じた形で自由に定義,記述することによって種々の機器や組み合わせに柔軟に対応できる。

FDML の基本構造を Table 1 に示す。情報の収集対象となる機器情報を記述する Info タグ,実フィールドでの物理的チャンネルと計測結果情報を管理する論理的チャンネルの関係を記述する Definition タグ,情報の取得・送信条件を記述する Condition タグ,当該取得データの時系列値を時刻情報と対応づけて記述する Data タグである。このように時系列データを構造化することによって,センシングで収集した人の情報,異なるタイプの情報の関係も統合的に扱うことができる。

Table 1. Main tags of FDML

Content of FDML	Description
<?xml version="1.0" encoding="Shift_JIS">	XML Declaration
<FDML version="1.02">	FDML start tag
<Info></Info>	Appliance information
<Definition></Definition>	Correlation with physical channels and logical channels
<Condition></Condition>	Condition to send FDML
<Data></Data>	Time series field-data
</FDML>	FDML end tag

#### 2.1.2. FDC

FDC (Field Data Center) プラットフォームサーバ

としては、FDML で構造化されたデータの蓄積・管理機能、FDML ファイルアクセスに対する認証・アクセス制御機能、FDML ファイルの監視・警報時処理機能を備えている。

またプラットフォームの提供機能として、基本的に収集した情報の蓄積・検索・閲覧とそれを利用した警報検知である加工・解析等の高度な機能は利用ケースに応じて情報の加工・解析手法も変わるので、一般的に汎用的なアプリケーションの提供は極めて困難である。そこで FDC プラットフォームサーバでは、蓄積された情報を活用した様々な外部アプリケーション連携を容易に実現するための、サーバ側で実行される組込 API、および外部連携 API をユーザに公開し、プログラマブルな仕組みを提供している。

この外部システム連携機能を利用することにより、ベンダや機器メーカーおよびユーザが、FDC サーバに蓄積された情報を利用し、個別の機器や工程、用途に応じた分析アプリケーションを開発することが可能となる。

```

<Definition>
<nChannel src="CSV2FDML">3</nChannel>
<Channel src="CSV2FDML" no="1">
<Name>瞳孔径</Name>
<Type>int</Type>
<Unit>pixel</Unit>
<Max>60</Max>
<Min>0</Min>
</Channel>
<Channel src="2FDML" no="2">
<Name>IBL</Name>
<Type>int</Type>
<Unit>ms</Unit>
<Max>15000</Max>
<Min>0</Min>
</Channel>
<Channel src="CSV2FDML" no="3">
<Name>心拍</Name>
<Type>int</Type>
<Unit>/min</Unit>
<Max>200</Max>
<Min>0</Min>
</Channel>
</Definition>
<Data>
<sample id="1">
<time>2004/01/16 11:32:18.859 JST</time>
<Group>QuizDailyChallenge_2004/01/16 11:32:18.859 JST_被験者B</Group>
<Channel src="CSV2FDML" no="1">
<value>40.90</value>
</Channel>
<Channel src="CSV2FDML" no="2">
<value>3267</value>
</Channel>
<Channel src="CSV2FDML" no="3">
<value>75</value>
</Channel>
</sample>
(略)
</Data>

```

Fig. 1. Example of Field Data Description using FDML

## 2.2. 設計コンセプト

今回開発した注目度推定プロトタイプシステムは FDC の外部連携アプリケーションとして実現した。FDC をベースとしたねらいは、次の 2 点である。

- 本注目度推定システムの要求する機能(映像コンテンツとセンシング情報の時間同期をとって蓄積・管理する、時刻に基づいて種々のデータを事後同期解析する)を FDC が既に備えていることから、

その機能を有効利用し、本プロトタイプ試作の開発期間短縮および開発コスト削減を図る。

- 一から開発するよりも、商用レベルの性能、品質を有している FDC の外部連携アプリケーションとして実装することにより、短期間で効率的に本プロトタイプの完成度を高める。

## 2.3. システム構成

視聴映像の注目度推定システム構成の一例を、Fig. 2 に示す。本システムでは、赤外線照明装置を備えた赤外線カメラと映像再生兼視線計測 PC によって視聴者の視線位置ならびに瞳孔径の情報を収集する(視線計測システム FreeGaze [7])。市販の DV カメラと瞬目計測 PC によって瞬目の頻度や時間間隔の情報を収集し、情報蓄積サーバに逐次蓄積する。

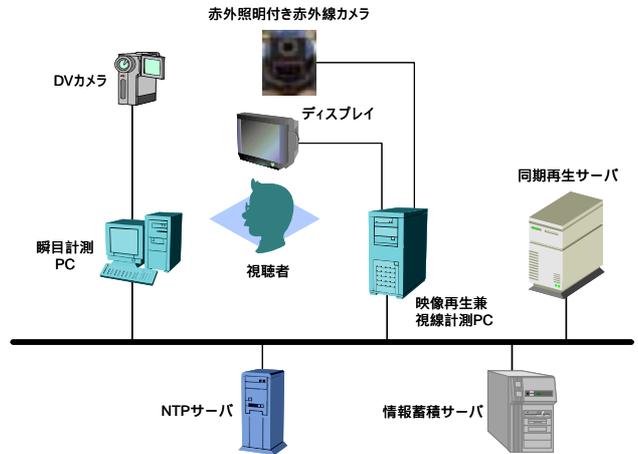


Fig. 2. Example of the attention-level estimation system configuration

## 2.4. 機能概要

### 2.4.1. 注目度推定機能

詳細は第 3 章で後述するが、現プロトタイプシステムでは初期段階として、瞬目時間間隔などの計測情報から視聴映像に対する注目度合いを推定する機能を実装した。将来的には、視線の移動パターンや瞳孔径の変化、頭部の動きなどのセンシング情報を統合することで、本推定機能の強化を図る。

### 2.4.2. 時間同期再生機能

詳細は第 5 章で後述するが、本システムでは前記情報蓄積サーバに蓄積された複数種類の時系列センシングデータを、マスタとする映像情報の収集開始時刻(または再生開始時刻)および再生時間から計算される時刻情報、ならびにデータ収集時に個別に付加された時刻情報に基づき、時間同期させて再生する機能を備える。

## 2.5. 同期再生の動作フロー

提案システムにおける時間同期再生機能の動作を、Fig. 3 を使って以降説明する。

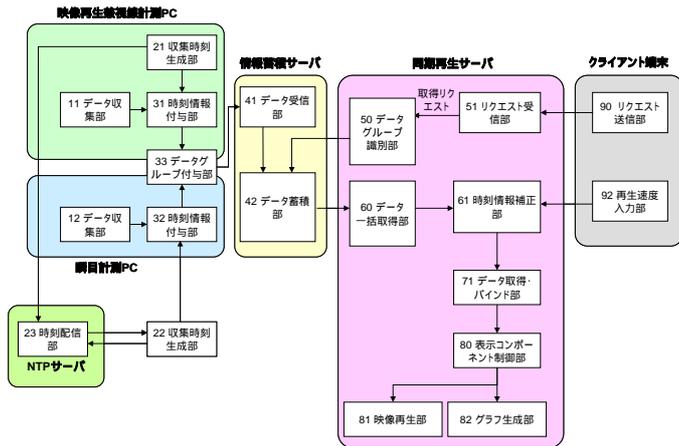


Fig. 3. Block diagram of our attention-level estimation system

Step1: 本システムでは、視聴者のリクエストに対して、映像再生兼視線計測 PC が同期再生サーバに蓄積された映像コンテンツファイルを取得・再生し、ディスプレイに表示する。赤外線カメラおよび映像再生兼視線計測 PC は、当該映像コンテンツ視聴時における視聴者の右眼あるいは左眼付近の画像を逐次キャプチャし、処理を行い、前記ディスプレイ画面上の視線位置および瞳孔径を自動計測する [7]。DV カメラおよび瞬目計測 PC は当該映像コンテンツ視聴時における視聴者の両眼付近の画像を逐次キャプチャし、処理を行い、瞬きの頻度や瞬目時間間隔などの情報を自動計測する。このように、映像再生兼視線計測 PC、および瞬目計測 PC におけるデータ収集部は、それぞれ一定周期で、視線の移動パターンや瞳孔径の変化、瞬目時間間隔などのセンシング情報をそれぞれ取得する。

Step2: 映像再生兼視線計測 PC、瞬目計測 PC における時刻情報付与部は、NTP サーバとの通信により、随時時刻合わせを行うと同時に、当該データの収集時刻を追加してデータグループ付与部に送信する。そして、データグループ付与部は、収集された各データ、ならびに収集時刻情報にデータグループ名を付与し、それらをひとまとめにして FDML 形式に変換後、情報蓄積サーバに送信する。

Step3: 情報蓄積サーバにおけるデータ受信部はデータグループ付与部から送信された上記の時系列データを受信し、データ蓄積部で蓄積する。

Step4: クライアント端末から同期再生サーバに時間同期再生リクエストが送信されると、データグループ識別部が条件として指定されたデータグループ名から同期再生する必要がある時系列データの集合を特定し、情報蓄積サーバから当該時系列データを一旦ダウンロードする。同時に再生すべき映像コンテンツファイルをバインドする。

Step5: 時刻情報補正部で収集手段ごとに時間

軸のずれがあれば補正する。

### 3. 提案する注目度推定法

#### 3.1. 注目度の定義

本研究における「注目度」とは、「(興味の有無に関わらず)実際に視聴者がどの程度集中して映像コンテンツをみているかという度合い」と定義する。

提案手法説明の前に、以下の用語定義を行う。

- 映像コンテンツ  
後述するリアクション情報、リアクション映像を取得するために人が視聴する映像
- リアクション情報  
視線位置情報、瞬目情報、心拍数など、カメラやセンサを用いて映像コンテンツを視聴している人から収集される情報の総称
- リアクション映像  
映像コンテンツ視聴時の視聴者を撮影した映像

#### 3.2. 注目度推定法の概要

本手法は人の負担が少ない非侵襲センシングデバイスである DV カメラを用いた画像センシングにより、視聴者の瞬目を検出し、瞬目時間間隔 (IBLI) 測度を算出する。そして、長い IBLI 測度とその後続く短い IBLI 測度の瞬目生起順序パターンに着目し、視聴映像に対する注目度を推定する。

### 4. 予備実験

#### 4.1. 実験概要

提案した注目度推定法の有効性を検証するために、ある情報バラエティ番組のクイズコーナー (長さは 13 分 32 秒、CF はカット) を題材として予備実験を行った。裸眼の被験者 6 名 (男性 5 名、女性 1 名) に対し、視聴中は出題されるクイズ全問に解答し、かつその正誤をメモするという課題を課した。

- 情報バラエティ番組のクイズコーナー
  - (a) 出演者紹介シーン (一部)
  - (b) クイズ出題・出演者解答シーン (一問分)
  - (c) 正解 VTR シーン (一問分)

#### 4.2. 注目度推定の結果

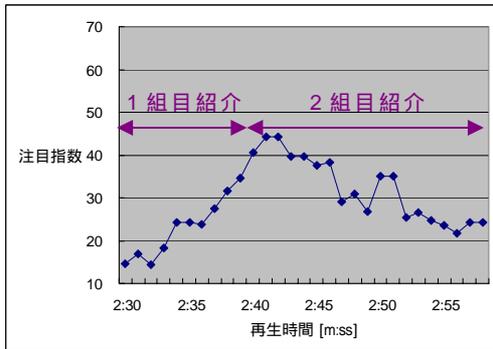
各シーン別に注目度の推定結果を Fig. 4 (a), (b), (c) にそれぞれ示す。折れ線グラフの横軸は視聴開始からのコンテンツ再生時間、縦軸は注目指数 (0 から 100 までの整数値) である。

(a) のシーンでは、最初の 10 秒で 1 組目、それ以降の 20 秒で 2 組目の解答者チームが紹介されている。(a) のグラフで 2 分 41 秒前後に注目指数が 44 となる一つのピークがみられるが、それはその変わり目とほぼ一致している。

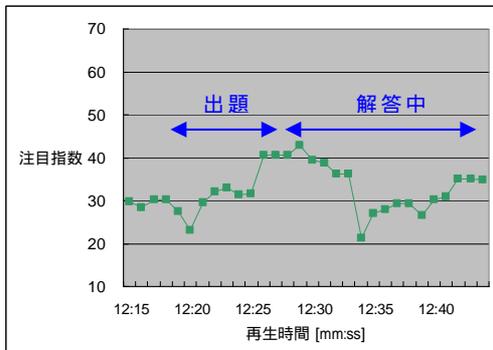
(b) のシーンでは、「インスタント袋めん つなげると一杯分の麺の長さはおよそ 40m である」は正しい

か否かというクイズが出題され、その後ゲスト回答者 3 組が約 10 秒間解答を考えて各自の解答を提示する。前節の課題設定により、被験者は映像を見ながらクイズの解答を考えているという状況が作りだされている。この状況下でも注目指数は 20 から 40 程度となっており、シーン (a) と大差はない。

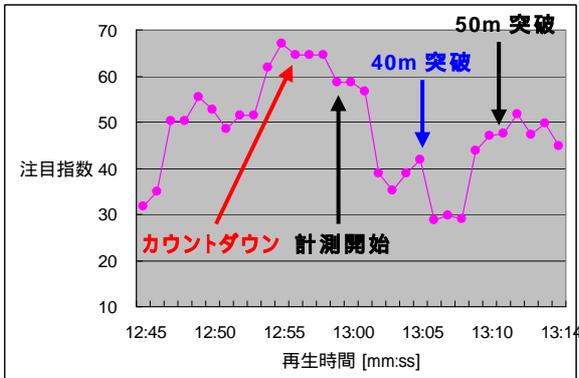
(c) のシーンでは、正解 VTR が流れて正答であることが判明する。VTR 前半、屋外で番組スタッフ数名が袋めん 1 本ずつをテーブルでつないでいるシーンのあと、実際に計測開始するまでのカウントダウン (3 秒前) の時点で注目指数のピークがたっている。番組演出の影響は無視できないが、正答と直接関係する 50m 突破の時点に比べても注目指数が高くなる結果が得られたことは非常に興味深い。



(a) Guest panelist introduction



(b) Setting question and thinking time



(c) Answer VTR

Fig. 4 (a)-(c). Experimental results of the attention-level estimation

## 5. 時間同期再生

### 5.1. データグループに基づくデータバインド

FDML/FDC では、データ収集時刻とデータの値の組それぞれに対して、当該データが属するデータグループ名を付与することで、収集蓄積した情報を統一的にハンドリングする。例えば、データグループ DG1 に属する時刻 Ta1 のデータが A1 であった場合、time タグには Ta1、value タグには A11 がセットされる。

情報蓄積サーバは同じデータグループに属する時系列データの集合を一つの単位として管理しており、このデータグループ名は一意である必要がある。データグループ名のフォーマットを、映像コンテンツのタイトル (日本語可) と当該映像コンテンツの収集開始時刻 (または再生開始時刻)、アンダーバー ( \_ ) を用いて、例えば以下のように規定する。なお、ユーザは映像コンテンツのファイルを同期再生サーバに登録する際に、当該映像コンテンツのタイトル名、ファイル名、およびファイルの保存場所 (同期再生サーバ側 / クライアント側)、その他付加的情報を入力、登録する。したがって、映像コンテンツの登録時に映像コンテンツのタイトルとそのファイル名がリンクされるので、データグループ名と映像コンテンツファイルは一意に対応づけられる。

**データグループ名のフォーマット: 映像コンテンツ**  
**タイトル名**\_\_YYYY/MM/DD hh:mm:ss.iii JST  
 ( 例 : QuizDailyChallenge\_\_2004/02/25  
 12:34:56.789 JST )

このようなデータグループ名規定によって、同期再生サーバはデータグループ名から時間同期させて再生・表示する必要がある時系列データの集合を特定し、情報蓄積サーバから自サーバに当該時系列データをすべてダウンロードする。

### 5.2. 実装方式

視聴映像の注目度推定システムの表示画面例を Fig. 5 に示す。図中の左上が映像コンテンツを再生するプライマリ映像表示コンポーネント、右上が視聴者のリアクション映像を再生する非プライマリ映像表示コンポーネント、下部は時系列センシングデータおよびそれら加工したデータを表示するグラフ表示コンポーネントである。この例では視聴コンテンツに対する注目指数を折れ線グラフとして表示している。

プライマリ映像表示コンポーネントの再生ボタンを押す、プログレスバーを動かすといった表示時刻を変化させるようなイベントを当該コンポーネントからその他のコンポーネントに通知する。その際、プライマリ映像表示コンポーネントで再生中映像のその時点での時刻を受信側のコンポーネントに渡す。そして、

あらかじめ規定したイベントリスナーインタフェースにしたがって各表示コンポーネントを作成することで、プライマリ映像表示コンポーネントからのイベントを受け取るコンポーネントの作成が可能となる。この実装部分は JMF (Java Media Framework), Java Applet を用いて実装した。

### 5.3. 応用事例と考察

実装した時間同期表示機能を用いてはじめて可能となる応用例を示すことで、本システムの有用性を説明する。

2.5 節で述べた動作フローにより、同期再生サーバに蓄積された視聴映像コンテンツ、DV カメラによって撮影、蓄積された視聴者のリアクション映像、映像再生兼視線計測 PC によって計測、蓄積された視聴者の視線位置ならびに瞳孔径の情報、瞬目計測 PC によって計測、蓄積された瞬目の頻度や時間間隔の情報を、後刻すべて時間同期させて再生することが可能となる。

Fig. 5 のように、視聴した映像コンテンツ、視聴者のリアクション映像、および視聴映像に対する注目度の遷移をクライアント端末の一画面におさめる。プライマリ映像表示コンポーネントのプログレスバーを動かすことにより、当該映像コンテンツのどのシーンを集中して見ていたかを、当該映像コンテンツと見比べながら一目で把握できる。この点が、提案する注目度推定システムのマルチメディアデータ時間同期再生機能の優れた特長の一つといえる。

なお JMF 仕様の制約から同期再生精度は 0.1 秒である。すなわち視聴コンテンツ映像、視聴者のリアクション映像、および収集蓄積されたセンシング情報の再生には最大 3 フレーム分のずれが生じる可能性がある。

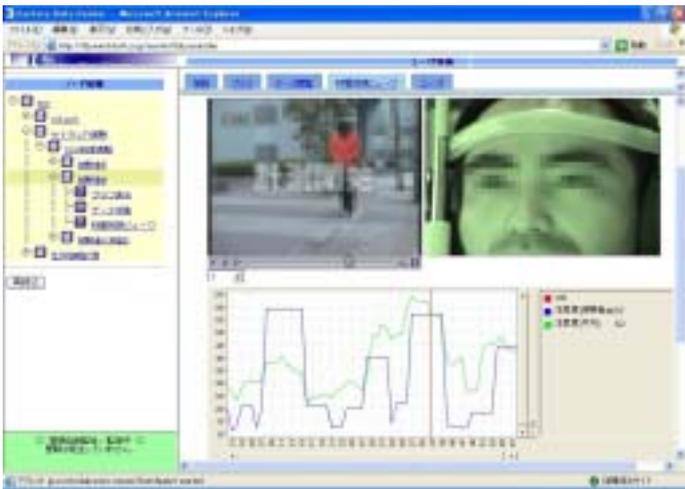


Fig. 5. Display screen example of the attention-level estimation system

## 6. むすび

本論文では製造業向け情報流通フォーマット FDML, およびフィールド情報流通プラットフォーム FDC を利用し、視聴映像に対する人の注目度推定に適用したプロトタイプシステムを提案・開発した。視聴映像コンテンツおよび、情報蓄積サーバに蓄積された視聴者のリアクション映像、視線位置ならびに瞳孔径の情報、瞬目の頻度や時間間隔の情報、心拍数の情報をすべて時間同期させて再生することによって事後同期解析が効率的に行えるという提案システムの有用性について論じた。

今後は被験者を増やした評価実験、同じジャンルの異なる映像コンテンツを視聴した評価実験を行うとともに、時間同期精度の定量的評価などを行っていく予定である。

## 謝辞

本研究を進めるにあたって、有益なコメント、ご助言を頂いた旧サービス基盤プロジェクト藤原進プロジェクトマネージャ、今枝尚史研究主任、グループ諸氏に感謝いたします。また、視線計測システム FreeGaze のプログラムを提供いただいた NTT コミュニケーション科学基礎研究所大野健彦研究主任に深く感謝します。

## 参考文献

- [1] T. Sato and T. Mori, "Realization of Robotic Room," Proceedings of the 3rd International Conference on Advanced Mechatronics, pp. 253-248, 1998.
- [2] H. Ishiguro, "Development of low-cost compact omnidirectional vision sensors and their applications", International Conference on Information systems, analysis synthesis, pp. 433-439, 1998.
- [3] 南, 杉田, 森川, 青山, "ユビキタス環境に向けたインターネットアプリケーションプラットフォーム", 信学会論文誌, Vol. J85-B, No. 12, pp. 2313-2330, 2002.
- [4] 清水, 岩田, 田野, "テレビ視聴時のマルチモーダル情報を利用した関心度推定システムの実現と評価", ヒューマンインタフェース学会論文誌, Vol. 5, No. 1, pp. 103-112, 2003.
- [5] 手塚, 遠藤, 今枝, 山本, 藪内, 下倉, "Field Data Markup Language (FDML) を活用した製造情報の構造化と情報流通", SICE SI2002 講演論文集(II), pp. 19-20, 2002.
- [6] 中山, 遠藤, 今枝, 山本, 藪内, 太田, 手塚, 浦野, 武藤, 茂木, 町野, 下倉, "製造業向け情報流通プラットフォームとその応用(遠隔監視・リモートエンジニアリング)", 電気学会産業応用部門大会(JIASC03), pp. II-83 - II-88, 2003.
- [7] Ohno, T., Mukawa, N., and Yoshikawa A. "FreeGaze: A gaze tracking system for everyday gaze interaction," Proceedings of Eye Tracking Research & Applications, pp. 125-132, 2002.