

# 差分画像を用いた定点カメラ映像分析支援インタフェース

野上 僚 司<sup>†</sup> 志築 文太郎<sup>†</sup>  
細部 博 史<sup>††</sup> 田中 二郎<sup>†</sup>

定点カメラによる人の行動分析は、撮影した映像が長期にわたるほど得られる情報は多くなる一方で、人間が全てを閲覧することは困難になる。本論文では、差分画像を用いて「変化のあった領域の提示」「変化のあった時間の提示」「類似映像検索」を実現することにより、長期の映像から、分析すべき映像を絞って閲覧することを支援するインタフェースを示す。

## An Analysis Support Interface using Frame Difference for a Video from a Stationary Camera

RYOJI NOGAMI,<sup>†</sup> BUNTAROU SHIZUKI,<sup>†</sup> HIROSHI HOSOBÉ<sup>††</sup>  
and JIRO TANAKA<sup>†</sup>

In analysing a long term video from a stationary camera, the longer we record a video, the more we get information. However, it is difficult to watch all in the long term video. In this paper, we present an interface supporting analysis of the long term video from the stationary camera. Using frame difference, the interface shows regions with changes, shows time when the changes occurred, and provides a video search capability based on similarity.

### 1. はじめに

定点カメラを用いた行動分析が行われている。実用例として、スーパーマーケット等での客層分析、動線分析等が挙げられ<sup>1)2)</sup>、また研究例としても広告の評価等に使われている<sup>3)</sup>。これらでは、システムが自動的にイベント（映像に映った人の数、人間の動線等）を収集し、システム利用者は取得されたイベント情報を閲覧することによって分析を行う。

一方、我々は定点カメラを用いた分析方法として人間が直接映像を閲覧して分析するという方法に着目する。人間が直接映像を閲覧して分析を行えば、「映像に映っている人は急いでいる様子だった」「ゆっくりしていた」等のコンピュータには取りだすことのできない情報を取得できる。しかし、映像を人間が閲覧して分析する場合、人間が映像を判断できる程度の速さで映像を再生する必要がある。この為、分析対象となる映像が数カ月に及ぶ場合この分析は現実的ではない。



図1 全方位カメラ画像

我々は、これまでにコミュニケーション支援を目的としたライブカメラ映像閲覧インタフェース<sup>4)</sup>を開発した。このインタフェースは、部屋の天井に設置した全方位カメラを用いて撮影した過去24時間のカメラ映像を素早く探索する為のインタフェースである。実際にカメラから撮影し得られた画像を図1に示す。

本研究ではこのインタフェースを発展させ、長期間の映像を対象とした人間が直接映像を閲覧して行う定点カメラ映像分析を支援するインタフェースを開発する。我々は、定点カメラ映像分析支援を行う為、差分画像に着目した。差分画像とは、定点カメラ映像を構成するフレーム（以降、定点カメラ画像）毎に、前フレームと比べ輝度値が変化したピクセルを求め、この輝度値が変化したピクセルを1、変化のなかったピ

<sup>†</sup> 筑波大学コンピュータサイエンス専攻  
Department of Computer Science, University of Tsukuba

<sup>††</sup> 国立情報学研究所  
National Institute of Informatics

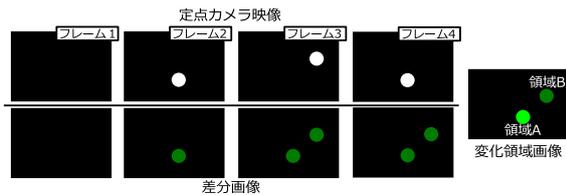


図 2 変化領域画像の例

クセルを 0 とした 2 値画像である。この差分画像を利用することにより、映像中変化のあった領域、すなわち人や物が動いた領域を求めることが可能である。この人や物が動いた領域を以降変化領域と呼ぶ。また、ある領域が変化領域となった時間を以降変化時間と呼ぶ。本インタフェースでは、「変化領域」と「変化時間」の提示、及び差分画像を利用した類似映像の検索を実現することにより定点カメラ映像分析を支援する。

## 2. 差分画像を用いた定点カメラ映像分析支援手法

我々は「変化領域の提示」「変化時間の提示」「類似映像検索」によって、定点カメラ映像分析を支援する。

### 2.1 変化領域の提示

分析の対象となっている期間（以降、分析対象期間）中、変化領域となったピクセル数（以降、変化量）が多かったピクセルほど明度の高い色で塗った画像（以降、変化領域画像）を提示する。ユーザは、変化領域画像を閲覧することにより期間中に変化が多かった領域、及び変化の少なかった領域を一目見て把握できる。図 2 に変化領域画像の例を示す。この例での定点カメラ映像は、4 フレームから構成される。この際、差分画像は 3 枚作成される。変化領域画像はこの 3 枚を加算合成することにより作成される。変化領域画像中の領域 A はフレーム 2, 3, 4 において変化領域となった領域であり、領域 B はフレーム 3, 4 において変化領域となった領域である。つまり、領域 A の方が変化量が多い領域である為領域 A の方が領域 B よりも明るい色で塗られる。この例では変化領域を示す色の色相は単色を用いているが、緑色、赤色、青色を合わせて用い、複数期間の変化領域の比較を行う事が可能である。

### 2.2 変化時間の提示

分析期間中のある領域にいつどれだけ変化があったのかを提示する。ユーザは、この提示により詳細に閲覧すべき映像を絞ることができ、余計な映像を閲覧する手間を省くことができる。

### 2.3 類似映像検索

「変化領域の提示」「変化時間の提示」を用いて閲

覧する時間を絞り、映像を閲覧した後に用いる。閲覧した映像の中に注目すべき行動を行っている映像があれば、その類似映像を抜き出すことにより注目した行動の原因をより深く分析できる。

この類似映像検索にも、差分画像を用いる。この理由は、人物の動きが類似する映像は対応する差分画像同士の類似性から判定することが可能だと考えたからである。まず、2 画像  $a$  及び  $b$  の類似度  $sim(a, b)$  を、画像に対応する差分画像を低解像度化した画像（以降、低解像度差分画像）間の類似度とした。すなわち  $sim(a, b)$  を求める手順は次の通りである。 $a$  及び  $b$  に対応する差分画像の低解像度差分画像を作る。次に、低解像度差分画像の各ピクセルの値を要素として持つベクトル  $\vec{a}$  と  $\vec{b}$  を求める。最終的にこの 2 つのベクトルから求まる  $cos\theta$  を  $sim(a, b)$  とした。

$$cos\theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

2 画像間の類似度を用いて、2 映像  $A$  及び  $B$  の類似度  $sim(A, B)$  を次のように定義した。ここで、 $a_i$  及び  $b_i (i = 1, \dots, n)$  は  $A$  及び  $B$  を構成する画像である。

$$sim(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{\vec{a}_i \cdot \vec{b}_i}{|\vec{a}_i| |\vec{b}_i|}$$

## 3. 定点カメラ映像分析支援インタフェース

前節に示した手法を用い、定点カメラ映像分析支援インタフェースを実装した。本インタフェースは期間指定部、カメラ映像部、時間部から構成される。外観を図 3 に示す。

### 3.1 期間指定部

分析対象期間の始まりと終わりを指定するテキストボックスを持つ。ユーザは両者の年、月、日、時、分を入力し Confirm ボタンを押す事により分析対象期間を指定する。Confirm ボタンが押されると、カメラ映像部に分析対象期間の変化領域画像が提示される。

### 3.2 カメラ映像部

定点カメラ画像上に分析対象期間の変化領域画像を重畳表示する。この際の定点カメラ画像は、ユーザが時間部を用いて指定した時刻の定点カメラ画像となる。なお、時刻が指定されていない状態では、誰もいない部屋を定点カメラで撮影した画像となる。変化領域画像における変化領域は緑色で描画する。この変化領域画像は、分析対象期間の差分画像を加算合成し、ピクセル値の高いピクセルほど、高い明度の緑色で描画することにより作成する。

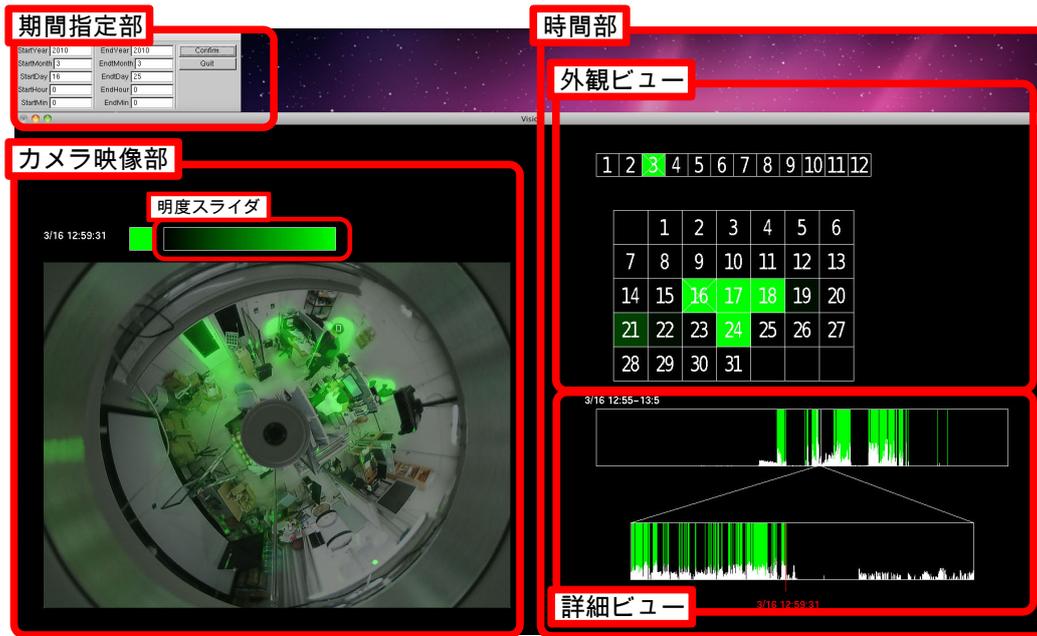


図 3 定点カメラ映像分析支援インタフェース

またカメラ映像部は、変化領域画像の緑色の明度を調節する為の明度スライダーを持つ。このスライダーは、変化領域の緑色の明度を決定するスライダー値を設定する。ユーザは、スライダーの領域内を左クリックした状態でドラッグすることによってスライダー値を変更できる。左にドラッグするほどスライダー値は低くなり、右にドラッグするほどスライダー値は高くなる。変化領域画像のそれぞれのピクセルの明度は、そのピクセルの変化量 × スライダー値で決定される。その為、スライダー値を低く設定し変化量の多いピクセルのみ人の目に見える明度を割り当てれば、最も変化のあった領域を見つけられる。その状態から、少しずつスライダー値を上げていけば、変化量の多くあった領域から順に緑色が見えてくる。スライダー値を操作することにより、変化領域の変化量の違いを、把握することができる。

また、ユーザは変化時間を求める為の領域選択にもカメラ映像部を用いる。画像内で左クリックした状態でドラッグすることによって、変化時間を提示する領域を矩形選択できる。この選択を行うと時間部に、選択された領域に対応する変化時間が提示される。

### 3.3 時間部

時間部は外観ビューと詳細ビューから構成される。2種類のビューによって、長期間の変化時間の外観を把握し変化時間の傾向を知る目的（外観ビュー）と、変化のあった瞬間を知り、実際の画像を閲覧する時刻を指定する、及び類似映像検索を行う映像を指定する目的（詳細ビュー）という異なる目的を実現する。

#### 3.3.1 外観ビュー

外観ビュー上部の1~12の数字は月を示す。それぞれの数字の背景色は、対応する月の間、かつ分析対象期間内の、カメラ映像部で選択された領域での変化量を示す。変化量が多ければ多い程、明度の高い色となる。この変化量は対応する期間の差分画像から求めることができる。ユーザはこの月の部分をクリックすることにより、その下方にクリックされた月の日を表示することができる。この日の背景色は月と同じく分析対象期間内の対応する日の変化量を示す。この日の部分をクリックすることにより、対応する日の変化時間を詳細ビューに提示することができる。

#### 3.3.2 詳細ビュー

詳細ビューは、外観ビューで指定された日の変化時間を提示する。1日分の変化時間を提示する上部のビューと、上部でクリックされた領域を拡大して10分間の変化時間を提示する下部のビューから構成される。上部のビューの横軸は時間軸であり、左端が外観ビューで選択された日の0時、右端が24時に対応する。1ピクセルは2分間を表し、1ピクセル毎に、対応する時間の変化量を色の明度によって示す線分が描画される。上部の提示がクリックされると、クリックされたピクセルの開始時刻から前5分間、後5分間の計10分間の変化量が下部に提示される。

下部のビューも、上部のビューと同じく横軸は時間軸であり、1ピクセルは1秒を表す。また、上部のビューと同じく対応する時刻の変化量を示す線分が描画される。この提示を左クリックすることにより、対応する

時刻の画像をカメラ映像部に表示することができる。また、この詳細ビューを用いて、類似映像検索を行うことが可能である。詳細ビューを右クリックした状態で、ドラッグすることにより、ドラッグした領域に対応する時間の映像と、詳細ビューに表示している日、1日分の映像との類似度が表示される。類似度は、上部のビュー及び下部のビューの白色の棒グラフによって表現される。棒が高い時間の映像ほど、指定した映像との類似度が高いことを示す。

#### 4. 実 装

本インタフェースを動作させる為のハードウェアとしてPCと全方位カメラ（シャープセミコンダクタ社製 LZ0P3551）を用いた。全方位カメラから得られる画像の解像度を 816×608、撮影頻度を 1fps とした。撮影した画像を PC に png 形式で圧縮して保存し続けている。2009年12月18日より撮影を開始し、現在も撮影中である。2010年12月16日現在約3200万枚を撮影した。本インタフェースの実装には開発言語としてC++を、ライブラリとしてOpenCV, OpenGL, GLUT, GLUIを使用した。

#### 5. 利用 例

本研究で想定している本インタフェースの利用例として、部屋に設置したキオスク端末がどのように利用されているかを分析する例を挙げる。前提として、分析を行う為に、キオスク端末を設置する1カ月前から定点カメラを設置して撮影を開始し、キオスク端末を設置してから1カ月経過した時点で撮影を止め、その期間中に撮影した映像を分析とする。

ユーザは、まず期間指定部によりキオスク端末設置前の1ヵ月間を指定し、カメラ映像部に变化領域を表示する。次に、キオスク端末設置後の期間を指定し、カメラ映像部に变化領域を表示する。この2つの变化領域を見比べることにより、キオスク端末設置後のみ変化が多くあった領域を見つけられる。ユーザは、この領域をカメラ映像部にて矩形選択し変化時間を表示する。

ユーザは時間部にて、変化のあった時間を左クリックにより指定し、実際の映像を閲覧する。人間がキオスク端末を利用している映像と、キオスク端末の前を通りすぎている映像を確認する。

ユーザは得たい映像がキオスク端末を利用している時の映像であると考え、キオスク端末を利用している時の映像を、右クリックにより指定して類似映像を検索する。類似度の高い時間を指定し、キオスク端末を

利用している映像のみを抜き出す。ユーザは、この映像を閲覧しキオスク端末の利用のされ方を分析する。直接映像を閲覧することにより、「端末利用者は片手で端末を操作しているのか、両手で端末を操作しているのか」「一人で端末を操作しているのか、友人と操作しているのか」等、コンピュータには判断しにくい情報を得られる。

#### 6. 関連 研究

橋本らは、ライブカメラより取得した30分の時空間画像をアルファブレンディングで合成し、その時間分の状況を一目で把握することを実現している<sup>5)</sup>。本研究は、期間のある映像を一目で把握することを目的としている点で関連するが、本研究は色を用いて過去の状況を表現する為、より長い期間の変化領域を提示可能な点が異なる。Romeroらは、定点カメラ映像の中で変化のあった領域とその量を提示する可視化を行い、撮りためたカメラ映像の分析を行う手法を提案している<sup>6)</sup>。変化領域の提示、変化時間の提示において関連するが、類似映像検索は行っていない。

#### 7. ま と め

本稿では、定点カメラ映像を分析する手法として、「変化領域の提示」「変化時間の提示」「類似映像検索」を提案し、提案した手法を用いたインタフェースを実装した。

#### 参 考 文 献

- 1) NEC Corporation: FieldAnalyst. <http://www.nec.co.jp/solution/video/fieldanalyst>.
- 2) Vitracom: SiteView. <http://www.vitracom.de/en/products.html>.
- 3) 南竹俊介, 高橋 伸, 田中二郎: 顔向き情報と移動軌跡を利用したデジタルサイネージの効果測定ツール, 情報処理学会第72回大会講演論文集, Vol.3, pp.323-324 (2010).
- 4) 野上僚司, 志築文太郎, 田中二郎: 過去の状況への気づきを支援するライブカメラ映像閲覧インタフェース, 情報処理学会シンポジウム インタラクシオン 2010 論文集, pp.181-184 (2010).
- 5) 橋本 悟, 中西泰人: ネットワークカメラ画像の合成による空間型コミュニケーションシステムの提案, 情報処理学会シンポジウム インタラクシオン 2006 論文集, pp.181-182 (2006).
- 6) Romero, M., Summet, J., Stasko, J. and Abowd, G.: Viz-A-Vis: Toward Visualizing Video through Computer Vision, *IEEE Transactions on Visualization and Computer Graphics*, Vol.14, No.6, pp.1261-1268 (2008).