

オフィス作業の"ぶった切り"ビデオ観察手法の提案 "Buttagiri": A method for a visual office work analysis

松村 耕平[†] 神田 陽治[†]
Kohei Matsumura Youji Kohda

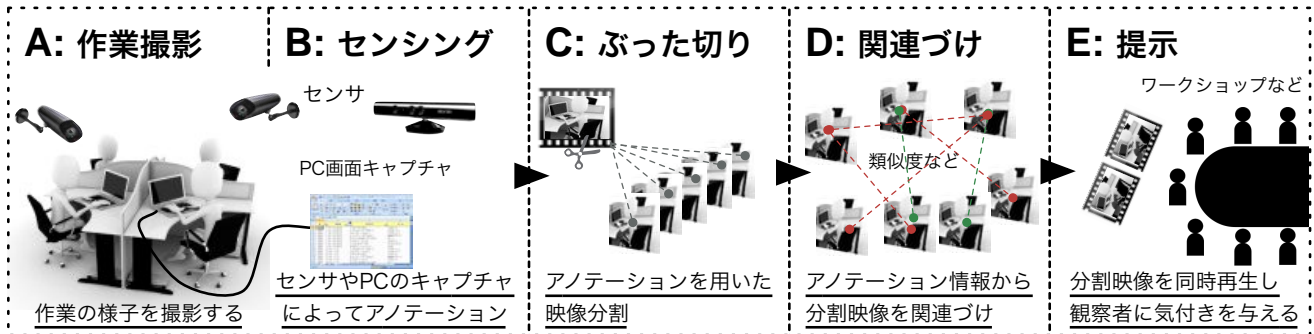


図1 ぶった切りフレームワーク

1. はじめに

オフィス等において、作業の効率化や質の向上のために、作業を観察・分析することが試みられている。

作業の観察・分析の方法としては、観察者が個人に張り付き、作業を目視・記録する、エスノグラフィ的な手法や、作業環境にカメラを設置し、全体的な動きや動線などを目視や画像解析により分析する手法がとられることが多い。前者は、個人単位での分析となり、チーム作業や、一人に閉じない作業（仕事が完了するために、複数人が直列に関わるようなタイプの作業）の分析や個人間の比較分析を行うことは難しい。後者は、工場などでの、作業の全体的な流れの把握、仕事の受け渡しでのボトルネックの発見が主な分析の対象となっており、オフィス作業の業務について分析・考察を与えることは困難である。

本研究では、オフィス作業において、作業の効率化や質の向上のために、観察者が気づきを得るためのビデオ観察手法について提案する。具体的には、オフィスなどに設置されたカメラ映像（作業ビデオ）を分割（ぶった切り）後、関連がある映像どうしを比較可能な形で、同時再生することにより、観察者が気づきを得る支援を目的とする。本論文では、ビデオ観察のためのフレームワークについて述べた後、開発中の“ぶった切り”の方法について説明する。

2. ぶった切りフレームワーク

2.1 ラピッドエスノグラフィ

従来行われてきた参与観察とインタビューによるエスノグラフィをより手軽に行う手法は、ラピッドエスノグラフィ[1][2]と呼ばれている。コンピュータ、センシングなどを活用した情報技術は、この手法について信頼できるデータの提供や時間の短縮に貢献できると考えられる。提案するフレームワークは、情報技術を用いて、ビデオ観察によ

るラピッドエスノグラフィを支援するものである。特に、人による参与観察ではコスト的制約から難しかったグループ作業の観察が、同時並行のビデオ撮影によって可能になると考え、研究を進めている。

2.2 フレームワークの構成

図1に、我々が提案するビデオ観察のためのフレームワークの概要を示す。このフレームワークは主に、

- ビデオによる作業の撮影
- A)によって撮影されたビデオについてアノテーションをつけるためのセンシング
- A), B)の解析による映像への機械的アノテーションによる映像分割
- 分割されたビデオ間の関係性の解析
- 観察者へのビデオの提示

の5つからなる。

A)では、作業員およびその作業環境を、ビデオカメラを用いて撮影する。オフィス作業ではPCは必須である。そこで、PC画面も作業環境に含める。

B)はA)で撮影されたビデオを分割するための情報を提供するセンシング環境である。特に、撮影されたビデオ自身もセンシング情報である。センサは環境に合わせて適切なものを選択する必要があると考える。例えば、Microsoft社からはKinectという奥行き情報が得られるカメラを用いた簡易なモーションキャプチャ装置が安価に売られている。Kinectからは、映像中の人物の数や動きなどのビデオ映像へのアノテーションに適した情報が得られる。

C)では、B)の情報或いはE)で観察者が入力したアノテーションによって映像を分割（ぶった切り）する。センシングされたデータの解析によって得られた作業の区切りといった高次の情報や、映像中に存在する人の数といった比較的低位（機械的に得られる）の情報が分割のために用いられる。

D)では、C)で分割された複数のビデオ間の関係性を計算する。C)で分割されたビデオには、分割のために用いたアノテーション及び、分割のためには用いられなかったアノテーションが含まれている。これらの双方の情報と、ビデオ

[†] 北陸先端科学技術大学院大学 知識科学研究科
School of Knowledge Science,
Japan Advanced Institute of Science and Technology

オの長さや記録された時間といったメタ情報を用いて複数ビデオ間の関係性を求める。

この関係性を元にして、E)で観察者へのビデオ提示を行う。ビデオ提示においては関係性を持つ複数の映像(≥2)を同時に再生することによって、観察者に気づきをもたらすように設計する。このビデオ提示中に、観察者によって手動でアノテーションを加えることが出来る。加えられたアノテーションは、C)やD)の処理のための情報として用いられる。また、観察者がアノテーション情報を元に動的にビデオを絞り込み・検索し、該当するビデオと、比較対象となるビデオを提示することなどの検索のためのインタフェースを備えることで、観察者が(検索によって)能動的に気づきを得ることを可能にする。

2.3 フレームワークの特徴

このフレームワークでは、「ぶった切り」によって分割された映像から、関連をもつ複数の映像を選択し、観察者に同時に再生して見せることにより、観察者が気づきを得ることを助けることを特徴とする。

スポーツや、職業トレーニングのためのビデオ教材の中には、複数例の比較から学習者に作業についての気づきを与えることを目指したものがある。同様に、オフィス作業においても、複数例を同時に再生することで、気づきを観察者にもたらしことができると考える。

一方で、提案するフレームワークに基づくシステム設計を行うためにはいくつかの課題がある。

3. 課題

提案フレームワークに関する課題として、§2で説明した5項目について考えるとA)では、撮影すべき対象はなにか。B)、C)に共通する課題として、ビデオの分割とビデオの再構成(比較ビデオの提示)のために適切なアノテーションはなにか。また、適切なアノテーションをつけるためのセンサはどのようなものか。D)では、分割されたビデオを関連付けるための方法はなにか。そして、E)では、観察者に対する適切な提示方法はなにか。など多くの課題があり、これらの課題を解決することが必要となる。

4. PC画面の解析によるビデオの分割

我々は、前述した課題「ビデオの分割とビデオの再構成のために適切なアノテーションはなにか。また、適切なアノテーションをつけるためのセンサはどのようなものか。」に対するひとつのアプローチとして、作業者のPC画面を録画し、その画面の特徴からビデオ分割のためのアノテーションを生成することを考えた。

画面特徴からビデオを分割する研究としては、例えばAhmetらによるサッカーのビデオを自動的に分割、要約する研究[3]や、番組のハイライトシーンを自動生成するものなどがあり、その手法は一定の適切さを持っていると言えるが、PC画面を解析し、ビデオを分割するためのアノテーションを生成するといった研究はされていない。

作業者のPC画面の特徴を作業ビデオへのアノテーションのために用いることの利点として、フィールドワークの際に簡便であることが挙げられる。PCの外部モニタ出力を録画することで、当該PCへのソフトウェアのインストールが不要となる。これは、現場への負担を軽減し、作業観察が受け入れられるためには重要である。

PC画面からの特徴抽出のためのアプローチは大きく2つある。一つは、時系列でのPC画面の特徴、すなわち画面そのものの変化量をみるもの。もう一つは、画面内に表示されている書類についての特徴をみるものである。

我々は、前者のアプローチのために、作業ウィンドウの変化パターンから作業の特徴を見るという方法を試している。具体的にはPC画面から、最も大きいウィンドウ様の矩形を検出することで、ウィンドウの抽出を行う。また、抽出された矩形について幾つかのピクセル間隔でグリッドに分割し、これらのグリッド内の輝度変化を抽出し、パターンを得る(図2)。例えば、文字入力中であれば、グリッドにより分割された画面は左から右に向けてその輝度が変化するパターンを、ページの移動は、画面全体の輝度が変化するパターンを持つ。

画面内に表示されている書類についての特徴抽出は、現時点では試していないが、例えば、OCRによって可能であると考えている。また、分割されたビデオの先頭および終端の映像中の文書をテンプレートとして記録し、ある一定時間における他の分割されたビデオの映像についてテンプレートマッチングを行うことで、作業の連続性などを検出できる。これは、§2の項目D)における関係性の発見のためのアプローチとして有効であろう。

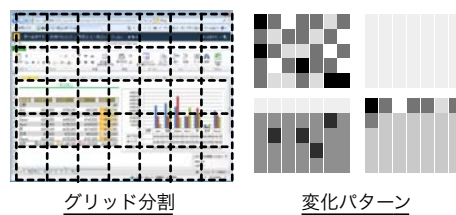


図2 PC画面の解析

5. まとめ

観察者にオフィスでのグループ作業についての気づきを与えることを目的とした、作業ビデオの再構成・提示のためのフレームワークを提案し、その実装のための課題について述べた。また、作業ビデオに対して、その分割のためのアノテーション付与の方法として、作業者のPC画面の特徴を用いる方法について示した。

提案したフレームワークでは、複数のオフィス作業の記録ビデオをいくつかのアノテーション手法を用いてフィルタリング操作を行うことで、動的に分割・再構成(ぶった切り)し、観察者に提示する。これにより、観察者がフィルタ情報と、それによって再構成されたビデオから気づきを得ることが期待される。

参考文献

- [1] David R. Millen. "Rapid ethnography: time deepening strategies for HCI field research", In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (ACM DIS '00), Daniel Boyarski and Wendy A. Kellogg (Eds.), 280-286, (2000).
- [2] 情報デザインフォーラム, "ユーザー調査のための手法", 情報デザインの教室, 第3章, 50-76, 丸善(2008).
- [3] Ekin Ahmet, Tekalp A. Murat and Mehrotra Rajiv, "Automatic soccer video analysis and summarization", *IEEE Transactions on Image Processing*, Vol. 12(7), 796-807 (2003).