

監視カメラ画像閲覧のための階層的画像集約手法

阿久津 渡 古谷 雅理 宮村 (中村) 浩子 斎藤 隆文
東京農工大学 大学院生物システム応用科学府

50005401202@st.tuat.ac.jp

監視カメラ画像において、管理者が人物の映りこむ画像を検索し閲覧することは非常に手間がかかる。本稿では、長時間撮影によって大量取得された監視カメラ画像を、1時間単位、または数分単位で集約し1枚の画像に表示する手法を提案する。従来手法のように、多岐の撮影条件の考慮、様々な閾値の設定をもとに、全自動で動物体を検知するのではなく、取得された全画像の情報を、何段階かの時間単位で集約し階層表示をする。これにより管理者に動物体の有無、動作を直感的に把握させ、効率的な閲覧を可能にする。

Hierarchical Image Gathering Technique for Browsing Surveillance Camera Images

Wataru AKUTSU, Tadasuke FURUYA
Hiroko Nakamura MIYAMURA, and Takafumi SAITO
Graduate School of Bio-Applications and Systems Engineering,
Tokyo University of Agriculture and Technology
50005401202@st.tuat.ac.jp

It requires large cost to inspect long time image sequences taken by a surveillance camera. We propose a new technique to generate a still image by gathering the moving parts from a image sequences by fixed camera. The gathered images are generated for several steps of time length, and hierarchically displayed. A user can easily browse the scene by looking at the images with moving parts. Since the detection and recognition of the target objects are done by human, efficient and reliable browsing can be established.

1 はじめに

本研究は、監視カメラで長時間撮影により大量取得された画像から集約画像を生成し、効率的に表示することを目的とする。

現在、監視カメラは防犯の事前事後対策において重要な役割を担っている。しかし、監視カメラは、常時撮影されるため取得される画像は膨大な枚数となる。そのため管理者が全ての画像を閲覧することは非常に手間がかかる。そこで、人物など動物体が映りこんだ画像をすばやく検索、閲覧できることが望まれる。人物など動物体を含んだ画像を全自動で検知することを目的とした研究があるが、それら従来の手法では監視カメラの多様な環境や条件により検出漏れ、誤認識を起こすことが問題とされる。

そこで本手法では、人物の検知や認識をせず、長

時間撮影された画像を、一定時間ごとに階層的に区切り、1枚の静止画像に情報を埋めこんで表示する。これにより検出漏れ、誤認識を生じることなく、全画像の中から管理者が閲覧すべき画像を容易に選択し、効率的に閲覧することができる。

2 動物体検知の関連研究

監視カメラの多くは、視野が固定されている。この場合、背景差分によって侵入者などの動物体の検出が可能である[1][2]。さらに、距離画像と濃淡画像を組み合わせたり[3]、複数個所のカメラからの画像を用いて検出する手法[4]。動物体の輪郭を追跡する研究がある[5][6]。

これらの従来手法は、動物体の自動認識を目指したものである。しかし、多くの場合、遠くに映った人物と、近くの小動物とを区別することや、照明条件、風などによる影響を排除することは困難であり、検出漏れや誤認識を完全に避けることは難しい。

3 集約画像

長時間撮影された画像を、任意の時間単位に区切って情報を圧縮し、1枚の集約画像に埋め込む。この画像を閲覧することで、管理者はその時間帯の動物体の有無を判断することができる。

3.1 集約画像の概要

Finkelstein らによって時空間解像度を必要に応じて階層的に変化させ、動画像を効率的に表示させる手法が提案されている[7]。

われわれは、Finkelstein らの時間解像度を変化させた詳細度制御の考え方を監視カメラ画像の閲覧に応用する。ただし、Finkelstein らの手法では、各画素を時間軸方向に平滑化した画像を用いているため、短時間だけ現れる動物体は、時間解像度を粗くするにつれて薄くなり消えてしまう。そこで、平滑化するのではなく、変化のあった部分を抽出して出力画像に埋めこむこととする。

本稿では、画像の集約する実時間のことを、集約画像の持つ画像集約時間とよぶ。画像集約時間を階層的に設定することによって、長時間撮影された大量画像の情報を詳細度制御する。

3.2 集約画像生成手順

集約画像は、画像集約時間内で背景から変化が生じた部分を寄せ集めたものである。したがって、画像をブロック単位に分割し、各ブロックごとに、フレーム間の輝度変化を調べ、変化が最大となるフレームを選択することで生成できる。図1に、画像集約時間内の画像から集約画像に埋め込む領域の選定方法を示す。

まず、各フレームを格子状のブロックに分割する。 n 番目フレームの注目ブロックにおける、輝度値の差分の平均値を以下の式で算出する。

$$L_n(Q) = \frac{\sum_{(x,y) \in Q} |B_{n+1}(x,y) - B_n(x,y)| + |B_n(x,y) - B_{n-1}(x,y)|}{2S}$$

$L_n(Q)$: n 番目フレームのブロック内輝度値変化量

Q : 注目ブロック

$B_n(x,y)$: 注目フレームにおける、画素 (x,y) の輝度値

S : ブロック Q の面積 全ブロック同一

次に、各ブロックに L_n を保存し、集約時間分のフレームで、同座標のブロックと比較する。座標ごとに L_n を最大にとるブロックのみ出力画像に表示する(図2)。これにより、任意時間の時系列画像における、輝度値の変化が大きく生じた時点の情報を、1枚の画像に集約し表示できる。

3.3 集約画像の生成結果

人物が撮影箇所を通過する画像に本手法を適用し、集約画像を生成した(図3)。出力されるブロックは、任意時間のフレームにおいて、 L_n が最大であったもののため、人物が重なる部分では、前後関係を無視して出力される。図3の太線で囲まれている部分のように人物が立ち止まった場合、集約画像では静止時間の情報が欠落する。しかし、通過する人物像が1箇所以上必ず含まれることから、人物の出現を見落とすことはない。

また、図4には、撮影箇所を通過した人数を数え、人数ごとに生成される集約画像の変化を示す。人数が増えていくにしたがい、画像は混雑する。通過した人数が非常に多い場合は、集約画像から細かな人物の動きをほとんど読み取る事はできない(図4(d))。しかしその場合でも、画像集約時間を短くした集約画像を生成すれば、詳細を読み取ることが可能となる。

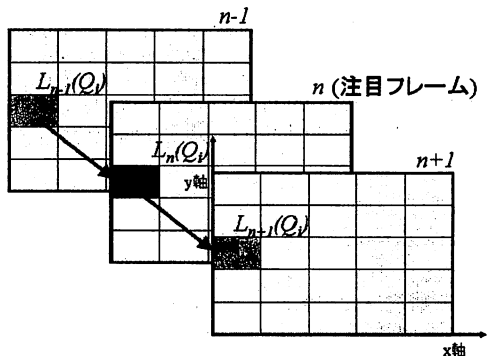


図1: L_n 算出における概略図

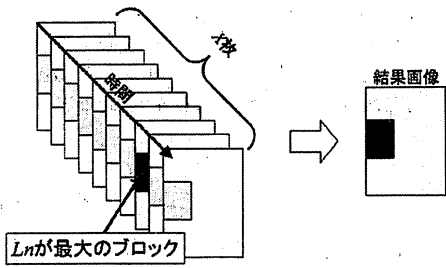


図 2: 出力画像の表示手順

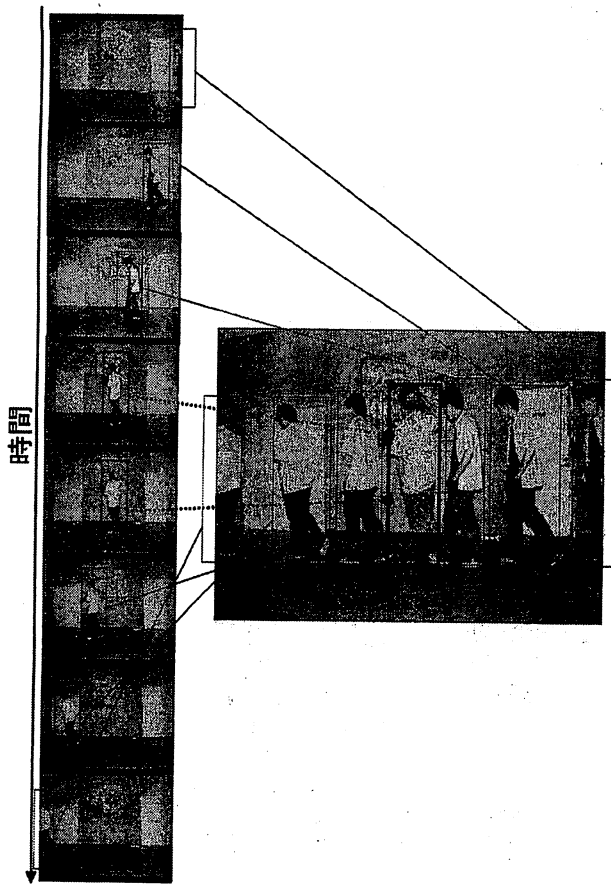


図 3: 集約画像の生成結果



(a) 1人



(b) 5人



(c) 10人



(d) 非常に多い場合

図 4: 通過した人数と集約画像結果

3.4 階層的閲覧表示

3.1 項で提案した集約画像を用いた、階層的閲覧表示を提案する。われわれは監視カメラ画像から、集約画像を異なる画像集約時間で生成し、それら集約画像を、画像集約時間をそろえて階層的に表示する(図5)。管理者は、画像集約時間の長い画像(例:1時間の集約画像)から閲覧し、そこで人物が映りこんでいるか否かを判断する。人物が確認できる集約画像に対して詳細を確認する場合は、下位階層の画像集約時間の短い画像(例:10分の集約画像)で追っていく。これにより、人物の映りこんだ画像だけを選択でき、効率的に閲覧することができる。

図6において例として、16時台に撮影された1時間分の監視カメラ画像に本手法を適用する。10分の画像集約時間を持つ集約画像を生成し、階層表示をすると、16時10分台と50分台に人物の出現が確認できる。更に、下位階層に1分の画像集約時間を持つ集約画像を階層表示することで、詳細情報を容易に追って見ることができる。また、本手法を用いることで、1時間分の画像を、わずかに数枚を閲覧するだけで確認できることがわかる。

とができない。しかし、人物の有無に関しては容易に確認ができる。これは管理者が人物の映りこんだ画像だけをすばやく確認し閲覧することを支援し、監視カメラで撮影された大量画像の効率的な閲覧を可能にする。

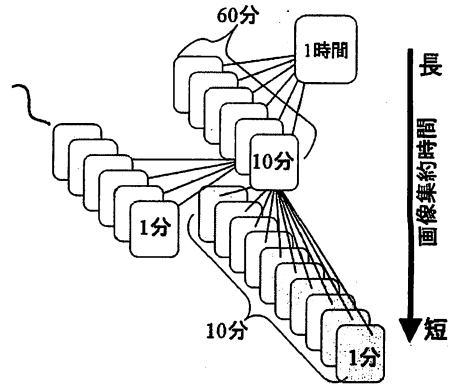


図5: 階層構造の概要

図7のように画像集約時間が長い場合は、人物が重なった画像になり、詳細情報を正確に読み取るこ

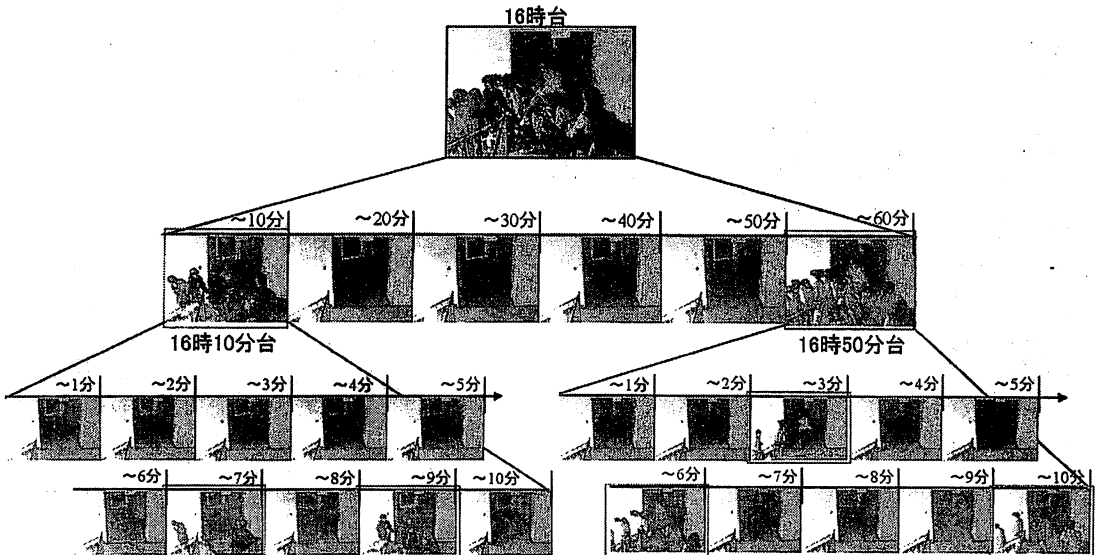


図6: 階層表示の例



図 7: 映りこみに置ける詳細情報の欠落

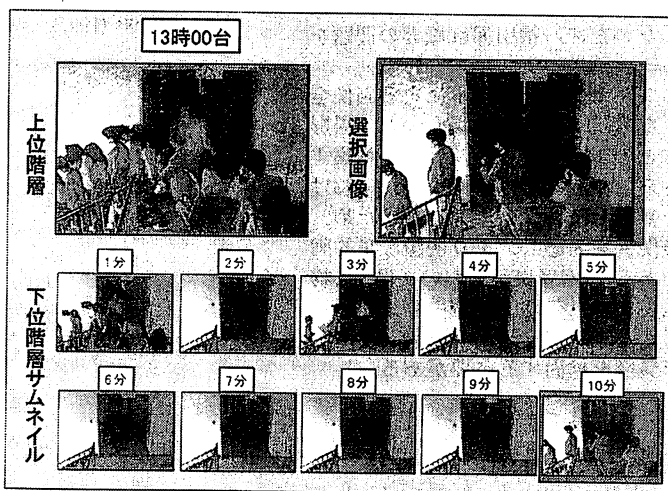


図 8: ツール概観 (画像集約時間 10 分)

3.5 階層的閲覧表示ツール

本稿では、容易に設置が可能で監視環境を整えられる Web カメラによる監視システムに適用し効率的閲覧を可能にするツールを作成する。

監視カメラの撮影環境

- ・ 固定カメラ
- ・ 画像解像度 704×480 pix
- ・ 撮影間隔 1 枚 / 秒

本手法の適用条件

- ・ ブロックサイズ 5×5 pix

実際の監視カメラ撮影で得られた大量画像に、階層的閲覧表示を適用する。管理者が実時間と対応付けをわかりやすくするため、24 時間を 1 時間、10 分、さらに 1 分の画像集約時間で、集約画像を生成して階層構造に構築する(図 8)。

実験により、人間は短時間で 10 枚程度の画像であれば、容易に全て閲覧できることから、一度に表示

する画像の枚数を設定し、効率的な閲覧を補助するツールを作成する。画像集約時間の短い集約画像を、図 9 の左上に表示させ、それらの下位階層における集約画像を、下部にサムネイルとして並べて表示する。これにより、上位階層の画像と下位階層の画像を照らし合わせ、比較しながら閲覧することができる。

並べて表示された、下位階層の集約画像を見ると、人物の映りこみがあるか否かを容易に判断できる。また、選択した下位階層の画像は右上に拡大表示される。

下位階層でも人物の詳細が十分に得られない場合は、確認したい下位階層の画像をクリックする。選択した画像が左上に拡大表示され、1 つ下の階層が下部にサムネイルで表示される。このように、人物の映りこんだ画像のみを選択、閲覧することで効率的な閲覧ができる。また、人物以外の動物体が映っていたとしても、管理者が閲覧の必要性を判断することができる。

4 評価と課題

集約画像においては、輝度値の変化が大きく生じた時点の情報を集約することで、時間変化をとまなう動物体の動作を1枚の画像に表示することができた。また、画像集約時間を変えた階層構造を持つ集約画像を生成することで、撮影された大量画像を効率的に閲覧することを可能にした。これは、時空間解像度の詳細度制御手法のコンセプトを応用しており、上位階層から動物体の有無など大局的な情報を得ると共に、下位階層から動物体の動作情報など詳細な情報を得ることができる。さらに本手法は、検出、認識は行なわないため、検出漏れなどの問題を解決した。

今後の課題として、センサーライトによる画像全体に輝度値の大きな変化が起こった場合など、撮影環境が生成画像にどのような影響を及ぼすか考察する必要がある。さらに集約画像は、撮影箇所における空間的な位置を無視して画像が出力されるため、それら物体の前後関係を考慮し、より明瞭な集約画像の生成が必要である。階層表示においては、人物が常に映りこむ環境では、効率的閲覧の効果を得ることが難しいと考え、本手法を適用し効果的である監視カメラの撮影環境を検討する必要がある。

5 おわりに

本稿では、監視カメラに映りこむ動物体の動作を1枚の画像に出力する集約画像の生成手法と、異なる画像集約時間を持つ集約画像の階層的表示により、監視カメラで撮影された大量画像を効率的閲覧する手法を実現した。

従来手法のように、検出や、認識などは用いていないため、検出漏れや、誤認識を生じない手法であり、画像生成は比較的簡単で高速に処理ができる。

参考文献

- [1] 波部 斉, 大矢 崇, 松山 隆司: “動的環境における頑健な背景差分の実現法,” 情報処理学会コンピュータビジョンとイメージメディア研究会画像の認識・理解シンポジウム MIRU'98, pp. 467-472, (1998)
- [2] S. Osher and J. A. Sethian: “Fronts Propagating with Curvature Dependent Speed,” *Algorithm based on Hamilton-Jacobi formation, J. Computational Physics*, Vol. 79, pp. 12-49, (1998).
- [3] 岡部 亜梨子, 小沢 慎治: “距離画像と濃淡画像を統合した侵入者検出システム,” 情報処理・産業システム情報化合同研究会, IP-04-18, (2004)
- [4] W. Hu, T. Tan, L. Wang, and S. Maybank: “A Survey on Visual Surveillance of Object Motion and Behaviors,” *IEEE Transaction on Cybernetics Part C: Application And Reviews*, Vol. 34, No. 3, (2004).
- [5] M. Kass, A. Witkim, and D. Terzopoulos: “Snakes, Active Contour Models,” *Int. J. Computer Vision*, Vol. 1.1, No. 4, pp. 321-331, (1988).
- [6] J. Sethian: “Level set Methods, 1st ed,” *Cambridge University Press*, New York, (1996).
- [7] A. Finkelstein, C. E. Jacobs, and D. H. Salesin: “Multiresolution Video,” *In Computer Graphics*, pp. 281-290, (1996).