

大規模映像監視の実現に向けた 映像センシング処理最適化アーキテクチャの提案

有熊 威† 岩松 洋介† 北野 貴稔†

NEC データサイエンス研究所†

近年、重大化する都市部重犯罪へ対応するために、映像解析技術により重要施設全域の監視カメラから犯罪の兆候を見つけ、犯罪を未然防止する試みが注目を集めている。しかし、犯罪の検知に使われる高度な映像解析は多くの計算機資源を必要とするため、多数カメラからの大規模映像センシングに必要なサーバ台数の増大が実適用の課題となっている。本論文では、メディア内容適応型効率化技術による映像センシング処理最適化アーキテクチャを提案し、提案アーキテクチャによるサーバ台数削減効果について報告する。

1. はじめに

近年、重大化する都市部での重大犯罪へ対応するため、顔認識などの高度な映像解析技術で監視映像をリアルタイムに解析する映像センシングにより、犯罪の兆候を発見し、未然に犯罪を防止することで、都市の安全を高める試みが注目を集めている[1]。

このような映像センシングを街頭監視のような多数カメラ環境へ適用するためには、高度な映像解析処理の効率的な実行が重要となる。映像解析は扱うデータサイズが大きく処理アルゴリズムも複雑なため、計算負荷が非常に高く、多数カメラ映像をリアルタイムに処理するには、大規模なシステムが必要なる。したがって、限られた計算資源で必要な性能要件を実現するための解析処理を効率化が重要となる。

本論文では、映像センシング処理を効率化する映像センシング処理最適化アーキテクチャを提案し、その有効性評価について報告する。

2. 映像センシング処理の実行効率化の課題

一般的な大規模計算と異なり、映像解析の効率化では解析負荷が映像内容に応じて大幅に変動することが課題となる。例えば、顔認証では映像フレーム画像中の顔数に比例して、特徴量抽出処理が発生する。解析可能な解像度の顔を10顔程度撮影できるFull HD画像の場合、処理負荷の変動幅は10倍近くなる。さらに、4Kや8K等の監視カメラ映像の高解像度化により一度に撮影可能な対象物数がFull HD比で約16倍に増加することで、負荷変動幅が増大することが予想され、変動幅の大きな負荷を効率的に扱うことが課題となってくる。

従来の映像解析システムでは、フレーム画像単位で処理を並列・分散処理することで、リアルタイム処

理を実現することが広く行われている。しかし、フレーム画像単位での処理では、監視映像の高解像度化の進展に伴い、下記の課題が顕在化してきている。

【課題1】 対象物増加によりリアルタイム処理が困難
フレーム画像単位での並列実行では、各画像を単スレッドで処理するため、画像内の対象物数が増えると処理時間が伸び、リアルタイム処理が困難になる。

【課題2】 分散遅延によりサーバ間資源共有が困難
フレーム画像単位での分散実行では、解析に必要な高品質画像(MB オーダ/画像)のデータ転送遅延が大きく、サーバ間での分散処理が困難である。

このように、大規模映像センシング処理を実現するためには、対象物増加による処理時間の増大と、分散オーバーヘッドの増加を抑えつつ、映像内容により大きく変動する負荷を、限られた計算資源で効率的に処理できることが求められる。

3. 映像センシング最適化アーキテクチャ

提案する映像センシング最適化アーキテクチャでは、内容適応型のリソース割当による対象物単位での処理並列と、ピーク負荷分散機構によるサーバ間での処理分散により、効率的なリアルタイム映像解析を実現する(図1)。

更に、提案アーキテクチャでは、映像解析をデータ入力、対象物検出、対象物からの特徴量抽出(特抽)、結果出力の大きく4つのフェーズへ分割し、それぞれの処理をプラグインとして拡張実装できるよ

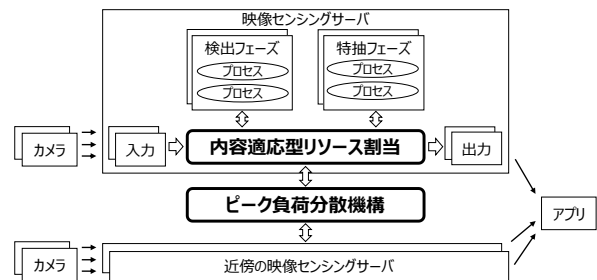


図1 映像センシング最適化アーキテクチャ

Proposal of sensing process optimization architecture for large-scale video analysis systems

†Data Research Laboratories, NEC Corporation

5-7-1 Shiba, Minato-ku, Tokyo, Japan

{t-arikuma@ce, y-iwamatsu@ab, t-kitano@ce}.jp.nec.com

設計した。これにより、顔や車、物体など多様な解析へリソース効率化機能を提供できる。

3.1. 内容負荷適応型リソース割当

負荷変動の主因が対象物毎の特抽処理にあることに注目し、解析処理を対象物の検出と特抽のフェーズに分割して、それぞれタスク割当を行う(表 1)。各カメラのフェーズ単位でプロセス数を増減可能となり、映像内容に合わせて最適な CPU コア数を割当てる事が出来る。さらに、各フェーズの解析プロセスがタスクを Pull することで、検出・特抽処理単体での負荷変動の影響を低減し、隙間なくタスクを割当てる事が出来る。このような並列化により、対象物数の変動があっても、フレーム処理時間を短縮し、限られた CPU 資源でのリアルタイム処理を可能とする。

表 1 各フェーズのタスク割当

フェーズ	タスク単位	割当	データ転送
検出	フレーム画像	Pull	共有メモリ
特抽	対象物画像	Pull	共有メモリ

3.2. ピーク負荷分散機構

各カメラの解析負荷ピークのずれに注目して、ピーク負荷のみを近傍サーバで分散処理する。このために、同一の解析機能を提供する解析サーバのうち、NW 的に近くに存在するサーバ群で、近傍サーバグループを構成する。各解析サーバは近傍サーバグループ内の他のサーバ群(N)の総体負荷(L)を下記式で算出し、自身の負荷(I)と比較して、Lの方が大きい場合は、対象物単位で解析タスクを分散タスクとして近傍サーバグループへ送信する。ここで、 l_i は各サーバの負荷、 h_i は処理性能などの重みである。

$$L = \frac{1}{n} \sum_{i \in N} l_i \cdot h_i \quad (1)$$

分散タスクは、近傍サーバグループ内で処理負荷が低いサーバが pull して処理する。このように、対象物単位で、必要最低限のタスクのみをサーバ間で分散実行することで、低オーバーヘッドの分散実行を実現し、資源効率を向上することができる。

4. 映像センシング最適化アーキテクチャの評価

4.1. 評価対象のシステムと映像

顔認証向けの映像センシングシステムを題材として、提案手法による資源効率化の効果を評価した。対象映像は海外の大規模主要駅の実監視映像を使用した。電車到着に応じた人の増減や店舗エリアのコンスタントな人の流れなど複数種の映像を用意した。

表 2 評価環境

CPU	Intel Xeon X5675 @3.07GHz x 2CPU
Memory	48GB
OS	Windows Server 2008R2
解析内容	顔検出・特徴量抽出

4.2. 評価結果と考察

4.2.1. 対象物単位での割当による処理遅延の短縮

提案手法とフレーム単位で並列化する従来でのフレームあたりの処理時間の分布を測定した(図 2)。秒間の顔数が 27 倍(0~27 顔/秒)変化する負荷変動が大きな映像において、最大フレーム処理時間を従来手法の 1,764msec から、提案手法では同 532msec まで最大 69.8%短縮できることを確認した。

4.2.2. ピーク負荷の分散実行による資源効率化

混雑ピークにより解析処理溢れが発生する映像で、提案手法とサーバ間負荷分散が無い従来手法で顔処理数を測定した(図 3)。Expected は十分な CPU コア数で溢れなく処理した場合である。従来手法では Expected を大きく下回る個所が発生しており、ピーク負荷時に処理溢れが発生している。提案手法は処理顔数が Expected とほぼ一致しており、サーバ間分散によりピーク負荷を処理できることを確認した。測定の結果、Expected に比べ最大 25.0%減の CPU コア数でリアルタイム処理が実現できることを確認した。

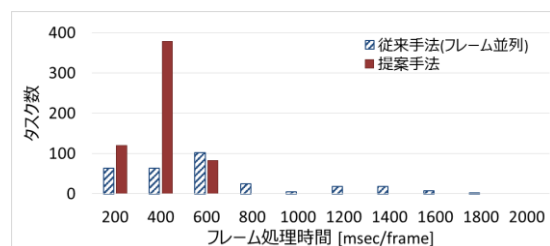


図 2 混雑映像におけるフレーム処理時間

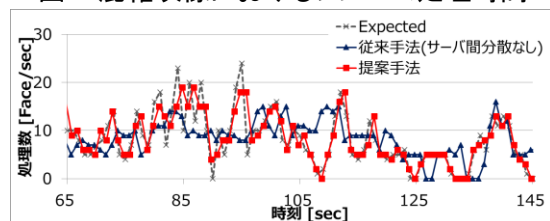


図 3 ピーク負荷時の処理顔数

5. まとめと今後の課題

高度な映像解析技術を活用して、監視カメラ映像から犯罪兆候などを発見する大規模映像センシングの実現に向けて、センシング処理最適化アーキテクチャを提案した。フレーム画像内の対象物毎にリソースを割当てる内容負荷適応型リソース割当により、対象物数の多い混雑映像でも、処理時間を最大 69.8%短縮し、リアルタイム処理を実現した。更に、ピーク負荷を対象物単位でサーバ間分散するピーク負荷分散機構により、最大 25.0%の解析リソース削減効果を確認した。今後は、フィールド検証を進め、アーキテクチャの有効性確認を進める。

参考文献

[1] P. Wnag, K. W. Woo, S. K. Koh, "Building a safer city in Singapore," NEC Technology Journal, pp. 71-74, 2015.