

大規模・高解像度解析を実現する映像内容適応型の負荷低減手法

Content-Adaptive Load Reduction Method for Large-Scale and High-Resolution Video Analysis

岩松洋介[†] 有熊威[†] 北野貴稔[†]YOSUKE IWAMATSU[†] TAKESHI ARIKUMA[†] TAKATOSHI KITANO[†]

1. はじめに

都市における犯罪防止に向け、重要施設に配置された監視カメラからの映像に対し多様な認識処理を実行することで、不審者の早期通知などの警備・業務支援を行う、映像解析システムの実現が求められている。

映像解析においては、近年の4Kなどの高解像度化に伴い処理負荷が増大しており、データの選別による負荷低減が必須となりつつある。しかし、撮影品質の低いデータを選択してしまうことによる解析精度の劣化が生じやすく、解析精度の維持とデータ削減率の向上の両立が課題となる。

本論文では、多様な解析に適用できる汎用性を持ちながら、映像内容に基づく高品質なデータの選別を可能にし、解析精度の維持と処理負荷の削減を両立する、映像内容適応型の負荷低減フレームワークを提案する。顔認識処理を対象とした、提案フレームワークの実解析への適用可能性の検証について報告する。

2. 映像解析における負荷低減の課題

2.1 解析精度の維持と負荷低減の両立

映像解析では、カメラの高解像度化により処理負荷が増大し、処理溢れなく全データを解析することが困難となりつつある。そのため、処理するデータを削減する必要があるが、解析精度を維持できるよう、有効なデータを選別することが課題となる。

高解像度化により処理負荷が増大する理由は、カメラ遠方の被写体でも識別が可能となり、映像から多数の被写体が検出されるようになるためである。被写体数が増加すると、それに比例して被写体毎の特徴抽出の処理数が増加し、解析に必要な処理負荷が増大するのである。

このように処理負荷が増大する中、全てのデータを解析することが困難な場合も多い。なぜならば、映像解析では、サーバ設置の制約等から、利用可能な計算資源が制限されることが多いからである。そのため、処理対象のデータを削減し、負荷を低減することが必要となる。

このような負荷低減の従来技術として、Tatbulらによるデータストリームの負荷低減手法がある[1]。この手法では、統計等から算出したデータ削減量と品質低下率の関係の予測を元に、品質への影響を抑えつつ削減するデータ量を最適化する。しかし、撮影品質が刻々と変化する映像解析の

ように、データの削減による品質低下の度合いが一定でないケースでは、意図せぬ解析精度の劣化が問題となる。

これに対して、筆者らの研究グループでは、これまで、映像内容に応じて負荷を制御可能なミドルウェアを提案した[2]。これを用いることで、無人時にフレーム画像を間引く等、映像内容に応じてデータを削減でき、解析精度を維持しながら負荷を低減することが可能となる。しかし、高解像度の映像のように、フレーム画像の中に撮影品質の異なる多数の被写体が映る場合においては、フレーム画像を単位としてデータを削減すると、品質の低い被写体のデータを選択し、解析精度が低下する恐れがある。

即ち、高解像度の映像解析システムにおいては、個々の被写体の品質といった映像内容の詳細な分析により、解析精度の維持とデータ削減率の向上を両立する、負荷低減処理が求められるのである。

2.2 多様な認識処理への対応

また、映像解析では、単一の認識処理ではなく、複数の認識処理を用いて映像シーンを分析することが必要となる。例えば、重要施設の監視では、顔認識を用いてリストにある不審者を発見し、発見した不審者を対象に、服色等の全身特徴を用いてカメラ間の移動を検知する等、多種の認識処理が求められる。

ここで、認識処理を実行する主体となるモジュールを、解析エンジンと呼ぶ。例えば、顔の検出、服色の特徴抽出を行うモジュールを、それぞれ顔検出エンジン、服色特抽出エンジンと呼ぶ。

映像解析システムの構築を効率化するためには、このような多種の解析エンジンに対する負荷低減処理に関し、個別の作りこみを極小化し、共通利用することが重要となる。

3. 映像内容適応型の負荷低減フレームワーク

本論文では、多様な解析エンジンに適用できる汎用性を持ちながら、映像中の個々の被写体の品質に基づいた有効なデータの選別を可能とする、映像内容適応型の負荷低減フレームワークを提案する。本フレームワークでは、解析処理を、フレーム画像からの被写体の検出、被写体毎の特徴抽出（以下、特抽と呼ぶ）の単位で分割することで、被写体単位でデータを選別可能な構成を用いる点を特徴とする。さらに、このような検出、特抽等の処理について、解析エンジン間で入出力データを共通化することで、処理の枠組みを多種の解析エンジンに適用可能とする。

[†] 日本電気株式会社, NEC Corporation

3.1 被写体毎の選別を可能とする検出・特抽分離型構成

提案フレームワークでのシステム構成を図 1 に示す。

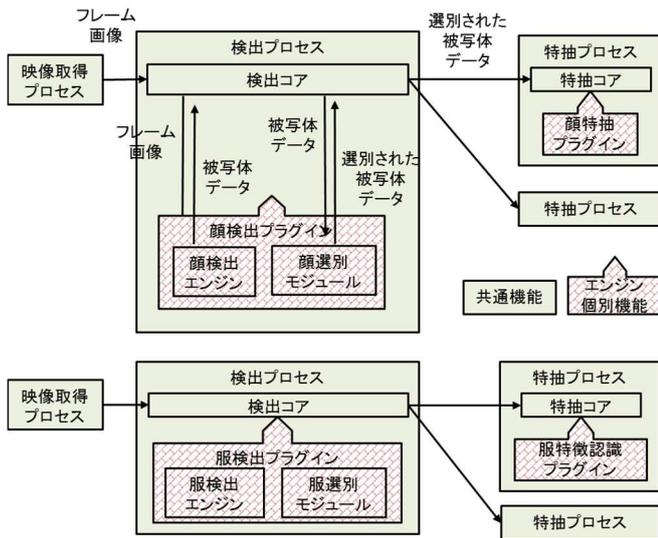


図 1 検出・特抽分離型のシステム構成

本フレームワークでは、フレーム画像からの被写体の検出と、被写体からの特徴抽出を、プロセスとして分離するアーキテクチャを採用する。分離した検出と特抽の間に、個々の被写体に対する選別処理を挿入可能である。この構成により、フレーム画像単位ではなく、個々の被写体単位での細やかな選別処理を可能とする。

3.2 多種の解析エンジンに対応する共通 IF

本フレームワークでは、検出や特抽処理のような解析の種別に応じて、多種の解析エンジン間で共通利用できる入出力モデルを提供する(表 1)。

表 1 処理種別と入出力

種別	入力	出力	概要
検出	画像	被写体データ	フレーム画像から被写体を検出。
選別	被写体データ	被写体データ	被写体を選別。
特抽	被写体データ	特徴量	被写体の特徴抽出。

例えば、顔の検出を行うエンジン、服の検出を行うエンジンについては、出力に用いる共通の被写体データを提供する。ここで、被写体データは、以下のフィールドを持つ。

- ・被写体の画像上の座標
- ・被写体のフレーム間の同一性を識別する追跡 ID
- ・被写体の撮影品質を示す品質スコア

このような共通の入出力モデルを用いることで、解析エンジン間で共通の機能と、解析エンジン個別の機能を分離することができ、多種の認識処理において共通機能を再利用することが可能である。本フレームワークでは、検出エンジン、データの選別モジュール、特抽エンジンを、認識処理に応じて拡張でき、その他の機能は共通的に利用可能である。これにより、顔の認識処理を服特徴の認識処理に置換する等、特定の解析に対して開発した負荷低減処理を、別の解析で容易に再利用可能とする。

4. 評価

4.1 評価方法

提案フレームワークを実解析に適用することで、負荷低減機能を有するシステムの構築が可能であることを確認するため、顔認識エンジンを対象に、提案フレームワークを用いたシステムを試作し、その負荷低減効果を評価した。

データ選別処理として、同じ追跡 ID 毎に被写体データを一定の時間幅でバッファリングし、最も品質の高いデータを選別する手法を用いた(図 2)。時間幅は 1.5 秒とした。

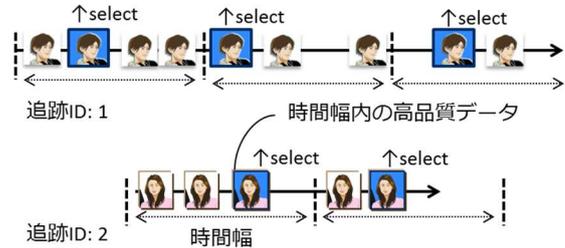


図 2 同一追跡 ID の被写体の選別

評価用データとして、屋外の通路や階段、海外の建物内エスカレータを含む 6 カメラからの実映像を使用した。

4.2 負荷低減効果

図 3 に、被写体データ削減効果を示す。被写体毎の選別処理により、顔認識処理において 64~77%の特抽の負荷低減が可能であることを確認した。

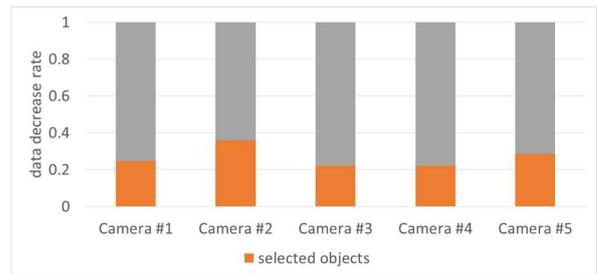


図 3 データ削減率

5. まとめ

高解像度の映像解析を実現するために、多様な解析に適用できる汎用性を持ちながら、映像内容に応じた高品質なデータを選別が可能で負荷低減フレームワークを提案した。提案フレームワークの適用により、顔認識の例において 64~77%など、大幅な負荷低減機能を有するシステムの構築が可能であることを確認した。

今後は、他の複数の解析エンジンへの適用と、解析エンジン毎のデータ選別アルゴリズムの効率化を行い、提案フレームワークの有効性検証を進める。

参考文献

- [1] Nesime Tatbul et al., “Load Shedding in a Data Stream Manager”, Proceedings of the 29th international conference on Very Large Data Bases, vol 29, pp. 309-320, 2003.
- [2] 有熊威ほか, “大規模映像解析システム向けの解析制御ミドルウェアの試作”, 第 11 回情報科学技術フォーラム, 4G-2, 2012.