Vol.2018-AVM-100 No.17
2018/3/7

鮮明化処理を考慮した高品質・低通信量な映像監視システムのための映像品質制御

坂牛和里[†1]　金井謙治[†1]　甲藤二郎[†1]　津田俊隆[†1]

概要：近年，Internet of Things(IoT)アプリケーションの一つとして映像監視システムが増加している．夜の時間帯の監視を想定した場合，映像の視認性が低下するため，鮮明化処理を行う必要がある．しかし，鮮明化映像は鮮明化前の映像よりも映像品質が劣化してしまうため，高ビットレートで圧縮する必要があり，映像トラフィックの増加の原因となる．本稿では，24 時間の監視の中で鮮明化処理を行う時間帯を適応的に制御し，監視対象でのイベントの検知結果により適切なビットレートを選択することで，冗長な映像トラフィックを削減する映像監視システムを提案する．また評価実験として，監視エリアを 24 時間撮影したシーケンスを利用し，適応制御の有無により人物検出精度や映像品質，映像トラフィックを比較することで，適応的な制御により監視の品質を維持しながら冗長な映像トラフィックを削減できることを示す．

キーワード：監視映像，画像鮮明化，レート制御，Internet of Things

Video Quality Control Considering Image Enhancement for High-quality and Network-friendly Video Surveillance System

Airi Sakaushi[†1]　Kenji Kanai[†1]　Jiro Katto[†1]　Toshitaka Tsuda[†1]

Abstract: Recently, video surveillance system is increasing as one of Internet of Things (IoT) applications. The video captured at night should be enhanced by using image enhancement techniques because the video has low visibility. However, the quality of enhanced video gets lower compared to non-enhanced video. Therefore, enhanced video should be encoded by high bitrate, but it causes increasing the video traffic. In this paper, we propose a video surveillance system that reduces redundant video traffic. The system controls the time of image enhancement processing for 24-hour surveillance and selects encoding rates adaptively. To validate the system performance, we capture the video at surveillance areas for 24 hours and evaluate the human detection accuracy, visual quality and video traffic. Results conclude that the proposed system can reduce redundant video traffic while sustaining surveillance quality by the event-driven bitrate adaptation.

Keywords: Video surveillance, Image enhancement, Rate adaptation, Internet of Things

1. はじめに

近年，IoT アプリケーションとして高解像度カメラや多様なセンサを活用した高度な映像監視システムの開発が進んでいる．このような映像監視システムが監視対象とする場所は様々であり，街中の道路や踏切，社会インフラ，自然環境などが挙げられる．このようなシステムでは，常に高品質な監視を維持すること，冗長なトラフィック，特に映像トラフィックを削減することが求められる．

高解像度及び高フレームレートのカメラを利用した映像を伝送することは，大量の映像トラフィックを生成することとなり，これによりネットワークの輻輳が発生しやすくなる．したがって，冗長なトラフィックを削減するためには，[1]にて筆者らが紹介しているような監視エリアの状態（イベント）に応じて動画像の圧縮レートを適応制御する必要がある．

また，24 時間の監視を想定した場合，時間帯により画像のコントラストは変化し，特に夜はコントラストが低く，視認性が低下する．そのため，視認性を高めるため，通常，コントラスト強調といった画像鮮明化処理(IE ： Image Enhancement)が広く利用されている．この処理は，コントラストが低い画像だけではなく，雨や雪といった悪天候により画像が不鮮明な場合にも利用される．

画像鮮明化処理による影響として，画質劣化の可能性が挙げられる．特に圧縮した映像の場合，コントラスト強調により圧縮時に発生したブロックノイズも強調されてしまうことが懸念される．このような鮮明化映像を，鮮明化処理を行っていない映像と同等の画質にするには，より高いビットレートで圧縮する必要がある．その一方で，高ビットレートは，映像トラフィックの増加に繋がるため，不必要な鮮明化処理は避ける必要がある．そのため，この鮮明化処理についても入力の画像特徴から適応制御することが求められる．

筆者らはこれまで，高品質な監視を維持しながら冗長なトラフィックを削減するため，[9]にて，圧縮映像に対する鮮明化処理を実施した画像において，画質評価，人物検知精度について評価してきた．さらにその結果を踏まえ，[7]にて，映像を圧縮するビットレートと鮮明化処理を適応制

†1 早稲田大学 基幹理工学研究科
Graduate school of Fundamental Science and Engineering, Waseda University.

ⓒ 2018 Information Processing Society of Japan

御する監視映像システムを提案している. 本システムでは, 入力された映像から監視エリアの状態検知と鮮明化処理判定を行い，それらの結果に応じて映像の圧縮レートを適応制御している.

本稿では，本提案システムを紹介するとともに，24 時間の監視映像を利用して，本提案システムの有効性を検証する. 具体的には, 24 時間監視映像のうち, 人物の検知状態, 検知状態および鮮明化処理判定により選択された画質，トラフィック量について定量評価する.

2. 関連研究

画像鮮明化処理は，画像のコントラストを持ち上げることで，暗い画像を明るい画像に変換する処理である．一般的な方法として輝度線形変換法とヒストグラム平均化法が挙げられる．輝度線形変換法は，原画像の最小輝度と最大輝度がそれぞれ 0 と 255 になるようにすべての輝度を線形に引き延ばす手法である．ヒストグラム平均化法は，一部に偏ったヒストグラムの分布を全体に一様に分布するように変換する手法である.

鮮明化処理を行う機器も多数開発されている．Infotech 社が開発した Red Super Eye G2[2]はその一つであり，リアルタイムに鮮明化処理を行うことができる．本機器では様々なパラメータを設定することができ，例えば，彩度，輝度，明度の強調度合いをそれぞれ調整するパラメータがある．本稿ではこの機器を利用して鮮明化処理を行っており，パラメータは，デフォルトで最適とされている値を設定する.

映像の圧縮符号化の品質評価やコントラスト感の評価指標も多数提案されている．圧縮符号化の品質評価には，PSNR (peak signal-to-noise ratio)や SSIM (structural similarity) [3]が広く利用される．PSNR を用いる画質評価手法は，原画像に対する純粋な画素値のズレを示しており，必ずしも人間の知覚と一致しているとは言えない．SSIM は人間の知覚とのずれを改善しており，画像の劣化が画像の構造の変化により知覚されると仮定している．VMAF (Video Multimethod Assessment Fusion) [4]も圧縮符号化の品質評価を行う評価指標の一つであり，Netflix により提案された機械学習ベースの画質評価指標である．VMAF の評価値は人間の知覚による主観評価値と高い相関があることでも知られている．本稿では，圧縮画像に対する鮮明化画像の画質品質を定量評価するため, PSNR および VMAF を採用する.

コントラスト感を評価する評価指標に関しては，Michelson コントラスト[5]が最も一般的な手法であり，画像の最大輝度の値と最小輝度の値の差を用いて表される．また，最小二乗誤差(RMS : Root Mean Square)コントラストは輝度の標準偏差をとることでコントラスト感を示している.

3. 提案する映像監視システム

3.1 目的

多様なセンサを用いた状態監視において，低ネットワーク負荷で効率的なセンシングを実現するために，筆者らは NICT による「ソーシャルビッグデータ利活用・基盤技術の研究開発[6]」に参画している.

これまで本研究課題において，筆者らは[1]にて，監視対象を，鉄道の踏切と想定し，踏切の安全向上のための映像監視システムを提案している．[1]では，高精度な人物検知を達成するために，画像ベースとセンサベースの両面で人物検知を実施している．さらに，検知結果に併せて映像符号化率を適応制御すること（イベントドリブン型レート制御）で, 冗長な映像トラフィックを削減している. 例えば, 人が踏切内に立ち止まっているときは，監視対象の危険性が高いと判断でき，高ビットレートで圧縮する．一方，人が踏切内を通過している場合あるいは無人の場合，危険性は低いと判断でき，映像は中あるいは低ビットレートで圧縮する．本監視システムは，エッジコンピューティングの活用を想定しており，カメラやセンサに物理的に近い場所に計算ノード（エッジサーバ）を設置し，エッジサーバで人物検知，映像品質制御，映像のキャッシュを実施している.

これら映像監視システムの拡張版として，先に述べたような画像鮮明化処理もエッジサーバで実施する．筆者らはこれまで，図 1 に示すように鮮明化処理が画質劣化を引き起こすことを示してきた[7]．図 1 は，非圧縮映像に対する鮮明化処理の有無による画像の圧縮特性を示している．PSNR の評価では，圧縮レートにかかわらず鮮明化映像の画質劣化度合いは大きいことを示しているが，主観評価との相関が高い VMAF においては, 圧縮レートが高くになるにつれ画像劣化度合いは小さくなる（＝画質劣化が目立たなくなる）という結果が得られている.

さらに，筆者らは鮮明化処理により人物検出精度の影響も評価している[7]．図 2 の評価結果が示しているように，昼のようなコントラストの高い映像に対して鮮明化処理を行った場合，人物検出の再現率の向上は確認できないが，夜のようなコントラストが低い映像に対して鮮明化処理を行うことで再現率を向上できることが確認できる.

以上のことから，コントラストが低い場合には，視認性を向上させるために鮮明化処理を行う必要があるが，鮮明化処理は画質劣化を引き起こすため，鮮明化処理を行っていない映像よりも高いビットレートで圧縮する必要があることがわかる．したがって冗長な映像トラフィックを削減するためには，鮮明化処理を適応制御しつつ，圧縮レートも適応制御する必要があると言える.

ⓒ 2018 Information Processing Society of Japan

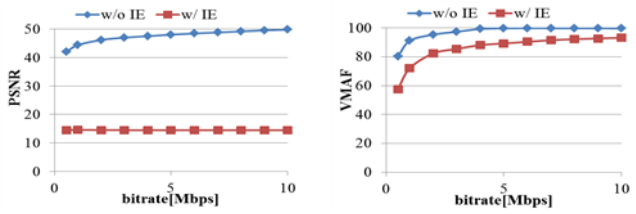


図 1　鮮明化処理の有無による圧縮特性の比較 [7]
(左：PSNR，右：VMAF)

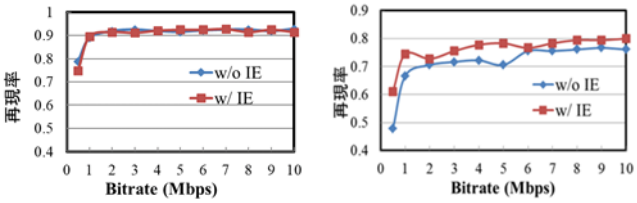


図 2.　鮮明化処理の有無による人物検出精度の比較 [7]
(左：昼の映像，右：夜の映像)

3.2 提案システムの概要

ここでは，筆者らが提案してきた高品質かつ低通信量なエッジサーバを活用した映像監視システムを紹介する[7]. 図 3 に提案システムの概念図を示す．提案システムは，(1)イベント検知，(2)鮮明化制御，(3)映像品質制御により構成されている．

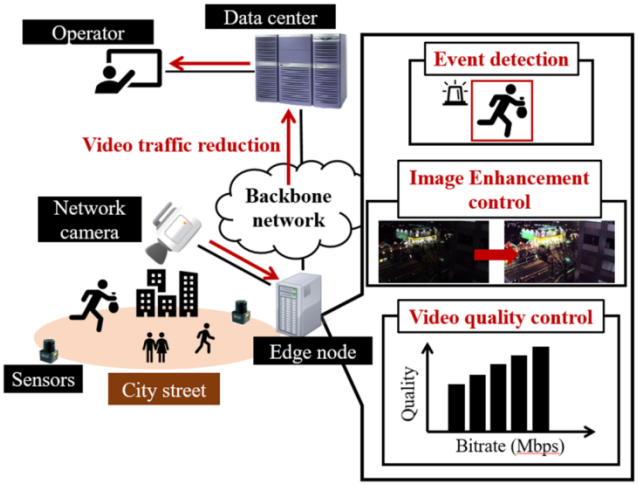


図 3.　提案システム [7]

まず，イベント検知では，入力画像やセンサの値を利用して管理エリアのイベント判定を行う．監視エリアのイベントを，例えば，図 4 に示すように，定常，注意，アラートの 3 状態に分類することとする. 次に, 鮮明化制御では，入力画像の画像特徴量によって鮮明化の有無を判定する．ここでは, 機械学習であるサポートベクターマシン（SVM）を用いて 2 クラス分類問題として扱う．学習させる特徴量には，画像の特徴量である YUV 表色系の輝度成分 Y の平均値，HSV 表色系の輝度成分 V の平均値，コントラスト感である RMS コントラスト，エントロピーの 4 つを利用することとする．最後の映像品質制御では，イベント検知による監視エリアの状態と鮮明化制御の結果により圧縮レートを決定する．例えば，図 4 の状態遷移図に従うと，定常状態が検知され，鮮明化処理を行っていない場合には低いビットレートで映像を圧縮し，一方，アラート状態が検知され，鮮明化処理を行う必要がある場合は高いビットレートで映像を圧縮することとする．この場合，合計 6 状態に対して，それぞれ圧縮レートを決定し，システムは選択されている状態に合わせて圧縮レートを決定することとする

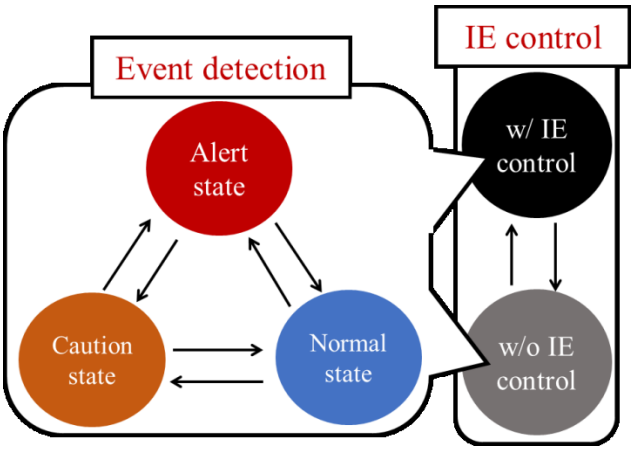


図 4. 提案システムにおける状態遷移の一例 [7]

4. 評価実験

本章では，3 章で述べた提案システムの評価について述べる．評価には，監視対象のエリアを 24 時間，晴れの日に 10 秒間ずつ 10 分おきに撮影したシーケンスを利用する．なお, シーケンスの解像度は 2K, フレームレートは 30fps, 非圧縮で記録しており，監視エリアは，研究室の窓から人通りの多い道路としている．なお人物検知には，[8]で提案されている深層学習ベースの検知アルゴリズムを利用した．

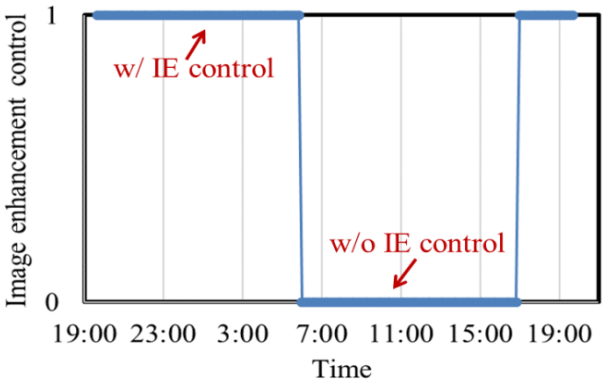


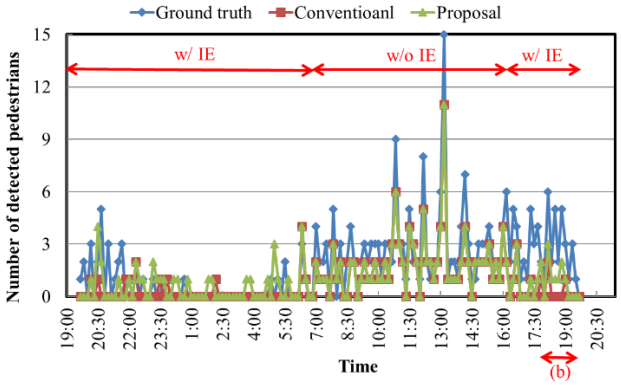
図 5.　24 時間における鮮明化制御の結果 [7]

4.1 SVM を利用した鮮明化制御結果

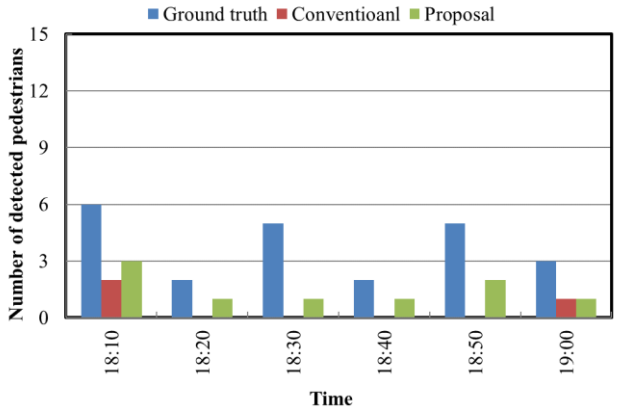
まず, SVM を利用した鮮明化制御の精度について評価す

る．学習データは鮮明化の有無が既知である 630 枚の画像とする．図 5 に鮮明化制御の結果を示す．図中，1 が鮮明化処理を必要と判断した時間，0 が不要と判断した時間と定義している．

図 5 より，6:00 頃と 17:00 頃が鮮明化処理判定の境界線となっていることが確認できる．この境界線の時間帯は日の出と日の入りの時間帯に近く，映像のコントラストが大きく変化していることから，また，目視による判断により SVM による鮮明化制御が適切に判定できていると確認した．



(a) 24 時間の全体の検知数推移



(b) 鮮明化処理が必要な時間帯の一例（18:10-19:00）

図 6. 24 時間における人物検知数の比較結果

4.2 人物検出精度評価結果

次に，提案システムによる鮮明化制御時と従来システムとして鮮明化制御なしによる人物検出数結果を比較する．人物検出対象はそれぞれの時間に撮影した映像の 1 フレーム目を利用している．図 6(a)に 24 時間全体の検知数推移を，図 6(b)に鮮明化処理が必要な時間帯である 18:10-19:00 の人物検出結果比較を抽出する．

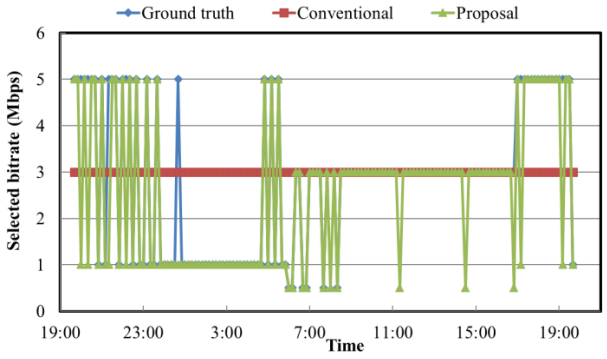
図 6 より，まず全体を見ると，いずれのシステムにおいても人物検知はおおよそ真値に従っているが，人数を正確に検知する点では，改善が必要であると言える．鮮明化制御の有無に注目してみると，鮮明化処理が必要な時間帯において，鮮明化処理を実施することで，人物の検出に成功していることが確認できる(図 6(b))．一方で，図 6(a)の深夜の時間帯(0:00～4:00)に注目すると，鮮明化処理を施すことで，誤検出が増えてしまうことも確認できる．検知精度向上については，映像のみならず他のセンサを活用することが挙げられるが，今後の課題とする．

4.3 映像品質評価

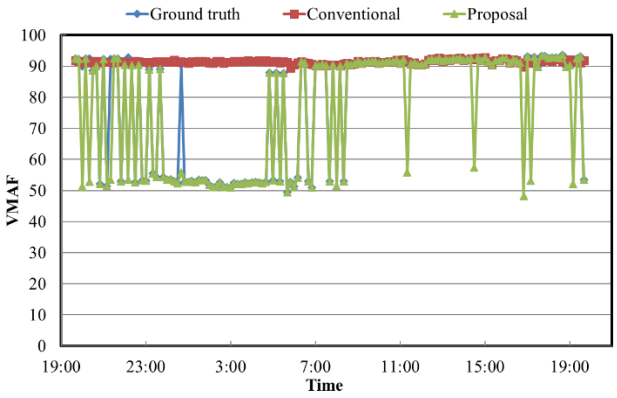
本節では，提案システムによって制御された 24 時間の画質の推移について評価する．単純化のため検知イベントは人の有無（＝2 状態）と定義し，人がいる場合を高品質状態，人がいない場合を低品質状態とする．それぞれの状態で選択されるビットレートは，表 1 のように定義する．これは，図 1 の結果を参考に，高品質の場合は，VMAF が 90 となるように，低品質の場合は，VMAF が 50 となるように設定している．なお，比較対象である従来システムでは，鮮明化制御なしで常に一定レート（3 Mbps）を選択するものとする．なお，真値（Ground truth）は，図 6 の結果に応じて，鮮明化処理も含め，適切にビットレートを選択したものと定義する．

表 1: 選択されるビットレート

	鮮明化有	鮮明化無
人がいる (高品質)	5 Mbps	3 Mbps
人がいない (低品質)	1 Mbps	0.5 Mbps



(a) 選択されたビットレート



(b) VMAF による画質評価

図 7. 24 時間シーケンスによる画質推移比較

ⓒ 2018 Information Processing Society of Japan

図 7 に選択されたビットレート推移と VMAF による画質の推移比較結果をそれぞれ示す．ビットレート制御は，前述の人物検知結果に強く依存していることから，誤検知，見落としの際に，不必要に画質を上下させてしまっている様子が見てとれる．一定レートを選択する場合，常に高品質になっている反面，冗長と言え，提案システムでは，監視の必要性が低い時間帯に画質を下げることに成功していることもわかる．

4.4 映像トラフィック量

最後に，これまでの実験結果を踏まえ，映像品質制御による映像トラフィック量の削減率を図 8 に示す．図より，イベント検知による適応レート制御を行うことで，映像トラフィック量を大きくは無いが，削減できるということが確認できる．今回，イベントの状態数を人物の有無の 2 状態と少なく，低品質状態が有効に機能しなかったことが原因であると言える．よりイベント状態数を増加させ，より柔軟に制御することで，トラフィックの削減効果を高める予定である．

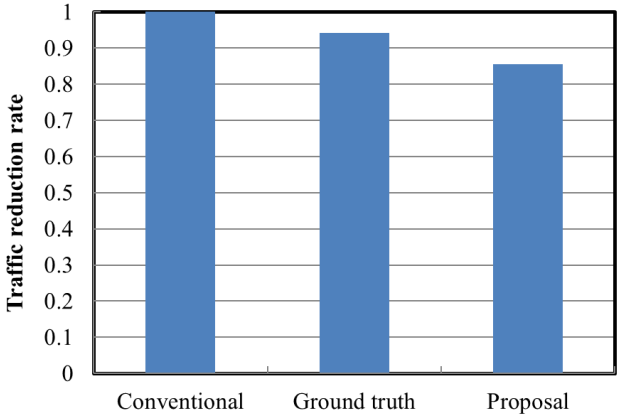


図 8　24 時間における映像トラフィック量の比較

5. おわりに

本研究では，高品質かつ低通信量な映像監視システムを提案した．常に高品質な監視を維持するため，映像のコントラストが低い場合には，画像鮮明化処理を適用することで視認性を向上させた．その一方で鮮明化映像が生成する映像トラフィックは大きいため，鮮明化処理が必要な時間は SVM により適応的に制御した．さらに，この鮮明化制御をイベント検知と組み合わせることで，人がいない時には映像を低ビットレートで圧縮することで冗長な映像トラフィックを削減できることを図った．24 時間の監視映像を利用した評価実験により，適応的な鮮明化制御ができること，映像品質制御を行うことで制御しない場合よりも映像トラフィックを削減できることを示した．

今後は，より高信頼なイベント検知のため，複数のセンサを組み合わせた手法を検討するとともに，より詳細なイベント状態を定義することで，より映像トラフィック削減を図っていく予定である．さらに，夜のシーケンスに限らず，雨や霧といった気象条件も考慮していく予定である．

謝辞　本研究成果は，JSPS 科研費 15H01684, 17K12681 の支援を受けている．

参考文献

[1] K. Kanai, K. Ogawa, M. Takeuchi, J. Katto and T. Tsuda, “Intelligent Video Surveillance System Based on Event Detection and Rate Adaptation by Using Multiple Sensors” IEICE Trans. Comm. Vol.E101-B, No.3, Mar. 2018. (to be published)
[2] Infotech: Red Super Eye G2 (Image enhancement equipment) [online]: http://www.infotech-japan.co.jp/res2_3ghd_s.htm.
[3] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structual Similarity,” IEEE Trans. Image Processing, Vol.13, No.4, pp.600-612, 2004.
[4] Z. Liu, et al, “Toward A Practical Perceptual Video Qaulity Metric,” Netflix Tech. Blog [online]: https://medium.com/netflix-techblog/.
[5] 瓦吹大，小野口一則，”ヒストグラム拡張とクラスタリングによる視界不良画像の鮮明化に関する研究，” 電気学会論文誌 C, Vol.136, No.10 pp.1473-1482, 2016
[6] NICT: “ソーシャルビッグデータ利活用・基盤技術の研究開発” [online]: https://www.nict.go.jp/collabo/commission/k_178b07.html
[7] A. Sakaushi, K. Kanai, J. Katto and T. Tsuda, “Edge-centric Video Surveillance System Based on Event-driven Rate Adaptation for 24-hour Monitoring,” PerFoT 2018, Mar. 2018. (accepted)
[8] OpenPose [online]: https://github.com/CMU-Perceptual-Computing-Lab/openpose.
[9] 坂牛，金井，甲藤，津田，“高品質・低通信量な監視映像システムのためのビットレート制御及び鮮明化制御，”情報処理学会 AVM 研究会, 2017 年 12 月．

ⓒ 2018 Information Processing Society of Japan