

観察者に応じたプライバシー保護映像を生成可能な映像配信手法

Delivery of Viewer-Specific Privacy Protected Video

福岡 直也[†]
Naoya Fukuoka

伊藤 義道[†]
Yoshimichi Ito

馬場口 登[†]
Noboru Babaguchi

1. まえがき

近年、治安の悪化を受けて監視カメラを用いた映像サーベイランスの普及が進んでいる。それに伴い防犯効果が期待される一方、「監視社会」といった言葉があるように、プライバシー侵害の問題も顕在化し始めている。このような状況において、単に監視社会を否定するのではなく、監視によって安全を確保し、同時にプライバシーを守ることに伴う映像サーベイランスに伴うプライバシーの問題を解決する方法が、近年数多く提案されている [1],[2],[3],[4],[5],[6],[7]。

その方法の一つとして、知野見らは、サーベイランス映像を見る者(観察者)と見られる者(被写体)との関係に応じて、透明化、ボックス表示、モザイク処理といった異なるプライバシー保護処理を被写体領域に施す手法を提案した [4],[5]。これにより、被写体それぞれが持つプライバシー感覚の差異を考慮した、柔軟なプライバシー保護を実現している。

このような観察者に応じたプライバシー保護映像を配信するための方法として、プライバシー保護処理を施すサーバがそれぞれの観察者の要求に応じてプライバシー保護映像を作成する方法が考えられる。しかし、この方法は、被写体と観察者に応じたプライバシー保護映像を全ての組み合わせに対して作成する必要があるため、サーバの負荷が非常に大きくなるという問題が生じる。サーバの負荷の観点からは、サーバが単一のプライバシー保護処理を施して映像を配信した後、観察者側が被写体との関係に応じてプライバシー保護映像を生成できるような方法が望ましいと考えられる。そこで本論文では、観察者側の端末で観察者の閲覧権限に応じたプライバシー保護映像を生成する映像配信手法を提案する。

これを実現するために、本論文では李らの提案した可逆型プライバシー保護画像処理の手法 [6],[7] に着目する。可逆型プライバシー保護画像処理とは、プライバシー保護処理の施された画像から元の画像を復元することが可能な画像処理である。Liらは離散ウェーブレット変換(Discrete Wavelet Transform: DWT)によって得られる低解像度成分を拡大することによってプライバシー保護画像を作成し、高解像度成分を情報ハイディングの手法を用いてプライバシー保護画像中に埋め込むことにより、可逆型プライバシー保護処理を実現している。本論文ではこの手法を応用し、被写体領域の情報をより基本的な情報に分解してそれらを映像中に埋め込み、こうして作成された映像から観察者が閲覧権限に応じてそれらの情報のいくつかを抽出して統合することにより、観察者に応じた様々なプライバシー保護映像を作成できることを示す。これにより、家族が閲覧する場合は実写表示、面

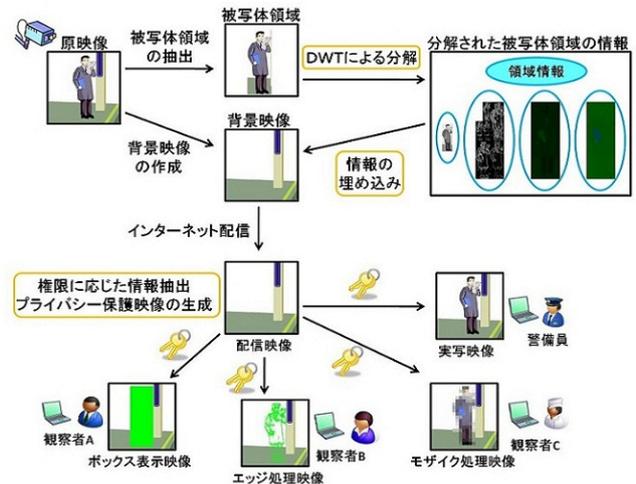


図 1: 提案手法の概要

識のない人物が閲覧する場合はボックス表示、といったように、被写体が持つプライバシー感覚に基づき観察者に応じて見せ方を変えることが可能となる。また、本手法は情報ハイディング手法を用いるため、被写体領域の情報を別途付加する手法と異なり、ファイルサイズの増加や同期に伴う問題が生じないという利点を有する。

本論文では、実験を通して埋め込む情報の量や映像の劣化の程度を調べ、提案手法を評価する。

2. 提案手法

提案手法の概要を図1に示す。まず、サーバは、サーベイランスカメラで撮影された入力映像から背景を推定して背景映像を作成するとともに、被写体領域を抽出する。次に、被写体領域に対しDWTを施すことによって得られる低解像度成分と高解像度成分を、被写体領域の位置情報及び色差成分とともに背景映像に埋め込むことにより、配信映像を作成する。配信映像を受け取った観察者側の端末では、埋め込まれた被写体領域に関する情報のいくつかをそれぞれの観察者の持つ権限に応じて抽出し、それらを統合することによってプライバシー保護映像を生成する。提案手法の主要部分は、次に示す3つの部分からなる。

1. 被写体領域の情報の分解
2. 分解した被写体領域の情報の埋め込みと抽出
3. プライバシー保護画像の生成

以下、これらについて詳しく述べる。

2.1 被写体領域の情報の分解

本節では、ボックス、エッジ、モザイクなどの様々なプライバシー保護画像を生成するために、被写体領域の

[†]大阪大学大学院工学研究科, Graduate School of Engineering, Osaka University

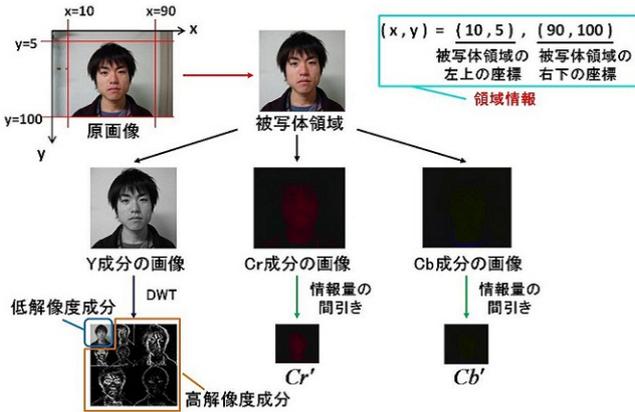


図 2: 被写体領域の情報の分解

情報を、領域情報、低解像度成分、高解像度成分、色差成分に分解する手法について述べる。被写体領域の情報の分解の流れを図 2 に示す。図 2 に示すように、まず、サーベイランスカメラから得られた入力画像に対し、ガウス混合モデル [8] を用いて背景モデルを作成し、被写体領域に相当する矩形領域を求める。ここでは、入力画像に対して xy 座標を設定する。このとき、被写体領域の左上を表す座標 (i_1, j_1) と右下を表す座標 (i_2, j_2) を領域情報と呼ぶ。図 2 の場合、領域情報は $(10, 5), (90, 100)$ である。

次に、被写体領域を (R,G,B) 表色系から (Y,Cr,Cb) 表色系に変換する。一般に、人間の目は色の变化より明るさの変化に敏感である。従って、輝度成分の劣化に対しては画質の劣化を感じやすいが、色差成分の劣化に対してはあまり画質の劣化を感じない。この特性を利用して、色差成分の情報量を $1/4$ に間引くことにより埋め込み量を削減する。入力画像のサイズを $2^M \times 2^N$ (pixel) とすると、色差成分の画像 Cr, Cb から情報量を間引くことにより得られる画像 Cr', Cb' は次式で与えられる。

$$Cr'(i, j) = Cr(2i, 2j) \quad (i = 0, \dots, M_1; j = 0, \dots, N_1),$$

$$Cb'(i, j) = Cb(2i, 2j) \quad (i = 0, \dots, M_1; j = 0, \dots, N_1).$$

但し、 $M_l = 2^{M-l} - 1, N_l = 2^{N-l} - 1$ である。

Y 成分に対しては、Haar の DWT を施し、低解像度成分 Y_s と高解像度成分 Y_w に分解する。 Y 成分の画像にレベル j の DWT を施すと、スケーリング係数 $S^{(j)}(p, q)$ および、水平方向、垂直方向、斜め方向のウェーブレット係数 $W_H^{(j)}(p, q), W_V^{(j)}(p, q), W_D^{(j)}(p, q)$ を得る。但し、 $p = 0, \dots, M_j; q = 0, \dots, N_j$ である。 Y_s, Y_w はそれぞれ、二次元配列 $S^{(j)}(p, q), W_H^{(j)}(p, q), W_V^{(j)}(p, q), W_D^{(j)}(p, q)$ から次式によって得られる一次元配列である。

$$Y_s = [S^{(j)}(0,0) \quad S^{(j)}(0,1) \quad \dots \quad S^{(j)}(M_k, N_k)],$$

$$Y_w = [Y_w^{(1)} \quad \dots \quad Y_w^{(j)}],$$

$$Y_w^{(k)} = [Y_{w,H}^{(k)} \quad Y_{w,V}^{(k)} \quad Y_{w,D}^{(k)}] \quad (k = 1, \dots, j),$$

$$Y_{w,X}^{(k)} = [W_X^{(k)}(0,0) \quad W_X^{(k)}(0,1) \quad \dots \quad W_X^{(k)}(M_k, N_k)]$$

$$(X \in \{H, V, D\}).$$

2.2 分解した被写体領域の情報の埋め込みと抽出

2.1 節で述べた、領域情報、低解像度成分、高解像度成分、色差成分に対し、適切なアクセス権を設けて情報ハイディング手法を用いて埋め込むことにより、観察者が抽出できる情報に制限を設けることができる。これにより、観察者の閲覧権限に応じたプライバシー保護画像が生成できる。本節では、プライバシー情報を埋め込むための情報ハイディング手法について述べる。

まず、二次元配列で表された色差成分 Cr', Cb' を、次式によって一次元配列 $\hat{C}r', \hat{C}b'$ に変換する。

$$\hat{C}r' = rs(Cr'), \quad \hat{C}b' = rs(Cb').$$

但し、 $rs(X)$ は行列 X の行展開であり、行列 $X = [x_{uv}]$ ($u = 1, \dots, m; v = 1, \dots, n$) に対し次式によって定義される。

$$rs(X) = [x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{m1}, \dots, x_{mn}].$$

また、領域情報 $(i_1, j_1), (i_2, j_2)$ を一次元配列 G による表現 $G = \{i_1, j_1, i_2, j_2\}$ に変換する。次に、一次元配列によって表現された低解像度成分 Y_s 、高解像度成分 Y_w 、色差成分 $\hat{C}r', \hat{C}b'$ を量子化幅 Δ で量子化する。なお、高解像度成分に関しては、0 周辺の値の出現頻度が高いため、0 に関するランレングス符号化を行い情報量を圧縮する。その後、これらと領域情報 G を 2 進数で表現し、データの区切りとして “2” を挿入することにより、配列 $E_G, E_{Y_w}, E_{Y_s}, E_{Cr}, E_{Cb}$ を作成する。さらにこれらの配列を、区切りを表す “22” を挿入して順に結合することにより、埋め込み用の配列 $E = \{e_1, e_2, \dots, e_s\}$ を作成する。以上により、配列 E は 0,1,2 から構成される配列となる。

分解した被写体領域の情報は、配信画像のビットマップ上に埋め込む。配列 $E = \{e_1, e_2, \dots, e_s\}$ の埋め込みには Amplitude Modulo Modulation (AMM) [9] を用いる。AMM による埋め込みは次式によって行われる。

$$\bar{F}_k = \operatorname{argmin}_{F \in S(e_k, 3)} |F - F_k|. \quad (1)$$

但し、 $F_k, \bar{F}_k (k = 1, \dots, 2^M \times 2^N)$ は、それぞれ k 番目の画素における埋め込み前後の画素値を表す。また、 $S(e_k, 3)$ は $S(e_k, 3) = \{F | F \equiv e_k \pmod{3}\}$ によって定義される集合である。式 (1) より、埋め込まれた値 e_k は、 \bar{F}_k を 3 で割った余りとして抽出できる。また、 R, G, B の 3 つの成分に埋め込むことを考慮すると、AMM によって配列 E を全て埋め込むためには、 $s \leq 3 \cdot 2^{M+N}$ である必要がある。

観察者は、自らの権限に応じて配信画像から埋め込まれた被写体領域の情報を抽出することにより、権限に応じた被写体領域のプライバシー保護画像を生成できる。生成手法は次節で詳しく説明する。

2.3 プライバシー保護画像の生成

提案手法において生成可能なプライバシー保護画像は、抽象度の高い順に、透明化、ボックス表示、エッジ処理、モザイク処理である。プライバシー保護画像生成に必要な情報を図 3 に示す。図 3 からわかるように、埋め込

	低解像度成分	高解像度成分	領域情報	色差成分
透明化画像	×	×	×	×
ボックス表示画像	×	×	○	×
エッジ処理画像	×	○	○	×
モザイク処理画像	○	×	○	○
実写画像	○	○	○	○

○: 必要 ×: 不要

図3: プライバシー保護画像生成に必要な情報

れた情報を抽出する権限を持たない観察者は、配信画像すなわち透明化画像しか閲覧できない。一方、全ての情報を抽出できる観察者には、実写を復元することが可能である。

まず、ボックス表示画像の生成法について述べる。ボックスは領域情報のみを用いて生成できる。抽出した領域情報 $(i_1, j_1), (i_2, j_2)$ に対し、これらをそれぞれ左上、右下とする矩形領域中の画素において、それらの画素値を $(R, G, B) = (0, 255, 0)$ に置き換えることにより、ボックス表示画像を生成できる。

次に、エッジ処理画像の生成法について述べる。エッジは高解像度成分、領域情報を用いて生成できる。まず、高解像度成分 Y_w の内、レベル1の高解像度成分に相当する $W_H^{(1)}(p, q), W_V^{(1)}(p, q), W_D^{(1)}(p, q)$ を重ね合わせた画像 $W_{HVD}(p, q)$ を次式により作成する。

$$W_{HVD}(p, q) = W_H^{(1)}(p, q) + W_V^{(1)}(p, q) + W_D^{(1)}(p, q). \quad (2)$$

次に、 $W_{HVD}(p, q)$ を2倍に拡大した画像 $W'_{HVD}(p, q)$ を作成する。そして、 $W'_{HVD}(p, q)$ の画素値が1以上であれば、対応する位置の配信画像の画素値を $(R, G, B) = (0, 255, 0)$ と置き換えることにより、エッジ処理画像を生成する。

続いて、モザイク処理画像の生成法について述べる。モザイクは低解像度成分、領域情報、色差成分を用いて生成できる。DWTのレベルが j の場合、低解像度成分に相当する $S^{(j)}(p, q)$ から成る画像を 2^j 倍に拡大することにより、モザイク処理の元となる画像の Y 成分を生成する。また、色差成分 Cr', Cb' を2倍に拡大することにより、 $\tilde{C}r, \tilde{C}b$ を生成する。それらを被写体領域の Y, Cr, Cb 成分の画像と置き換え、 (Y, Cr, Cb) 表色系から (R, G, B) 表色系に変換することにより、被写体領域の実写を生成し、領域情報を用いて配信画像中の対応する領域に上書きすることにより、モザイク処理画像を生成する。

最後に、実写画像の生成法について述べる。実写は低解像度成分、高解像度成分、領域情報、色差成分を用いて生成できる。低解像度成分、高解像度成分に対して逆DWTを施すことにより、被写体領域の Y 成分を復元する。 Y 成分と拡大した色差成分 Cr', Cb' を、 (Y, Cr, Cb) 表色系から (R, G, B) 表色系に変換することにより、被写体領域の実写を生成し、領域情報を用いて配信画像中の対応する領域に上書きすることにより、実写画像を生成する。

3. 実験

3.1 実験概要

本節では、撮影した映像に対して提案手法を適用し、量子化幅 Δ が、埋め込む情報の量や復元された実写映像の画質に与える影響を検討する。なお、埋め込む情報の量を表す尺度として、1画素当たりの埋め込み量を表す平均プライバシー情報量 (Average of Privacy Information: API) を用いる。ただし、APIは次式で定義される。

$$API = \frac{\text{埋め込み用の配列 } E \text{ の長さ}}{\text{被写体領域の総画素数}}$$

また、復元された実写映像の画質の劣化を表す尺度として、PSNR (Peak Signal to Noise Ratio) を用いる。実験に用いた映像について、画像サイズは 720×480 (pixel)、フレームレートは3フレーム/秒である。離散ウェーブレット変換のレベルは4とする。

3.2 実験結果

撮影された映像に対し、提案手法による処理を行った。結果を図4、図5に示す。図4は原映像と $\Delta = 1$ の場合の配信映像であり、図5は $\Delta = 1$ 及び $\Delta = 9$ の場合のプライバシー保護映像である。図5より、エッジ処理映像は量子化幅 Δ の増加によるエッジの劣化への影響が大きいことが分かる。これは、エッジの生成に用いる高解像度成分は0周辺の値の出現頻度が高いため、量子化幅を大きくすると高解像度成分が0となり、エッジが生成できなくなるためである。また、エッジ処理映像と比較して、モザイク処理映像とボックス表示映像は量子化幅 Δ の増加による画質の劣化への影響が小さいことが分かる。

また、量子化幅 Δ が、埋め込む情報の量や復元された実写映像の画質に与える影響を調べた。結果を図6に示す。図6より、量子化幅が大きくなるにつれて、埋め込むべき情報の量が少なくなることが分かる。 $\Delta = 1$ の時、APIは6.02である。埋め込むべき情報の量が最も少ない $\Delta = 10$ の場合では、APIは2.74である。提案手法では、画素の RGB 成分に配列 E の要素を1個ずつ埋め込むことによって、1画素当たり3個の要素を埋め込んでいる。よって、提案手法では $\Delta \geq 9$ であれば、被写体領域に全ての被写体領域の情報を埋め込むことができる。しかし、量子化幅 Δ を大きくした場合、画質の劣化が問題となる。図6より、量子化幅 Δ が大きくなるにつれて、被写体領域における原映像と実写映像のPSNRは低くなり、画質が劣化することが分かる。 $\Delta \leq 5$ の場合にはPSNRが30(dB)を超えているため画質の劣化が小さいが、 $\Delta \geq 6$ の場合、画質の劣化が大きいことが分かる。

また、 $\Delta \leq 8$ の場合、被写体領域の大きさによっては、被写体領域の情報の一部を失う可能性がある。現在の手法では、1画素当たり3個の要素の埋め込みを行い、画像中の全ての画素に情報を埋め込むと、その時点で埋め込み処理を終了する。そのため、全ての被写体領域の情報を埋め込む前に処理が終了した場合、残りの情報は失われ、生成したプライバシー保護映像の色合いに乱れが生じる。

現段階では、 $\Delta = 1$ の場合、被写体領域が映像全体の約50%を超えると、被写体領域の情報の埋め込みの途中



図 4: 原映像と配信映像

 $\Delta=1$  $\Delta=9$ 

ボックス表示映像 エッジ処理映像 モザイク処理映像 実写映像

図 5: 生成したプライバシー保護映像

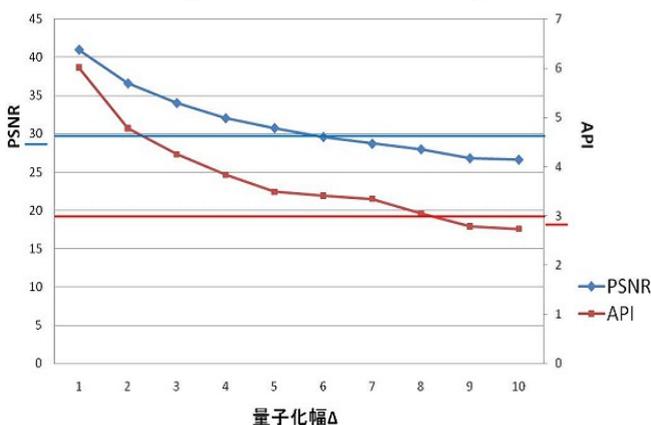


図 6: 量子化幅 に対する API と PSNR

で埋め込みが終了することにより、情報の一部が失われ、生成されるプライバシー保護映像に乱れが生じる。量子化幅 Δ を大きくすることによって情報を圧縮し、埋め込み量を減らすことができるが、 $\Delta \geq 6$ の場合、PSNR が 30dB を下回り、被写体領域の情報の劣化が問題となる。図 6 より、被写体領域の大きさや量子化幅に関わらず、被写体領域に関する情報を全て埋め込むためには、1 画素当たり 7 個程度の要素の埋め込みが必要であることが分かる。

4. まとめ

本稿では、観察者の権限に応じたプライバシー保護映像を生成可能な映像配信手法を提案した。分解した被写体領域の情報に対し適切なアクセス権を設けて配信映像中に埋め込むことにより、観察者は閲覧権限に応じて被写体領域の情報の一部を抽出し、モザイク、エッジ、ボックス、実写といったプライバシー保護映像を生成できることを確認した。また、撮影した映像に対して提案手法を適用し、量子化幅 Δ が、埋め込む情報の量や復元さ

れた実写映像の画質に与える影響を調べた。その結果、量子化幅の増加につれて埋め込む情報の量が減少し、それに伴い復元された実写映像の画質が劣化することが分かった。今後の課題として、埋め込む情報量の問題を解決するために、色差成分の間引きや 0 についてのランレングス符号化によって埋め込むプライバシー情報量を削減するだけでなく、より埋め込む情報量が小さくなるような埋め込みビット配列の作成手法を開発することが挙げられる。また、配列 E の要素を全て埋め込むために、より埋め込む情報の量が大きい埋め込み手法を検討することも課題となる。なお、本研究の一部は科学研究費補助金(基盤研究(A))による。

参考文献

- [1] F. Dufaux, and T. Ebrahimi, "Scrambling for video surveillance with privacy," *In Proc. of IEEE Computer Vision and Pattern Recognition Workshop on Privacy Research in Vision*, 2006.
- [2] W. Zhang, S.S. Cheung, and M. Chen, "Hiding privacy information in video surveillance system," *In Proc. of International Conference on Image Processing*, 2005.
- [3] X. Yu and N. Babaguchi, "Privacy preserving: hiding a face in a face," *In Proc. of the 8th Asian Conference on Computer Vision*, pp. 651-661, 2007.
- [4] 知野見 健太, 李 光鎮, 中 嶋大介, 新田 直子, 伊藤 義道, 馬場口 登, "PriSur: プライバシー保護機能を有する映像サーベイランスシステム," *情報処理学会論文誌コンピュータビジョンとイメージメディア (CVIM22)*, vol. 1, no. 2, pp. 152-162, 2008.
- [5] K. Chinomi, N. Nitta, Y. Ito, and N. Babaguchi, "PriSur: Privacy Protected Video Surveillance System Using Adaptive Visual Abstraction," *Proc. of the 14th International MultiMedia Modeling Conference*, pp. 144-154, 2008.
- [6] 李 光鎮, 伊藤 義道, 于 小億, 新田 直子, 馬場口 登, "可逆型プライバシー保護画像処理," *電子情報通信学会第 6 回マルチメディア情報ハイディング研究会*, DS-3-3, pp.S-23-S-24, 2009.
- [7] G. Li, Y. Ito, X. Yu, N. Nitta, and N. Babaguchi, "Recoverable Privacy Protection for Video Content Distribution," *EURASIP Journal on Information Security*, vol. 2009, article ID 293031, online published, 2009.
- [8] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection," *Proc. of the 2nd European Workshop on Advanced Video Based Surveillance Systems*, 2001.
- [9] M. Wu and B. Liu, *Multimedia Data Hiding*, Springer, 2002.