

非通常動作検出によるリスク発見を支援する 手術映像保存システムの提案

坂部 史生^{†1} 村川 正宏^{†2} 小林 匠^{†2}
樋口 哲也^{†2} 大津 展之^{†2}

手術室の状況を多面的に記録した映像を、リスクの発見に利用する取り組みが始まったばかりである。これを普及させるうえで、事後の映像チェックにかかる多大な手間と保存すべき映像のデータ量増加は解決すべき課題である。そこで本稿では、手術映像中の出現頻度が低い動作にリスクが潜んでいると仮定し、我々が提案したシーン適応型動画圧縮法を適用することで、この2つの問題を同時に解決する。具体的には頻度が高い通常動作から逸脱する動作シーンほど高品質に、その他のシーンを低品質に保存することで、手術映像全体のデータ量を削減する。さらに高品質に保存されたシーンにしおりを付与することで、事後チェックの手間を大幅に低減する。8.5時間の手術映像を用いた実験の結果、データ量を約68.3%削減することに成功した。

Proposal of Surgery Video Archiving System for Supporting Risk Detection Using Infrequent Motion Recognition

FUMIO SAKABE,^{†1} MASAHIRO MURAKAWA,^{†2}
TAKUMI KOBAYASHI,^{†2} TETSUYA HIGUCHI^{†2}
and NOBUYUKI OTSU^{†2}

In this paper, we propose a surgery video archiving system which supports to find risks in long recorded surgery video. Assuming that the potential risks lie in the motions that infrequently occur, the system reduces the video data size according to a coding efficiency which is proportional to frequency of motion in the recorded data. Moreover, chapter marks are placed just before the positions of the infrequent motions to reduce reviewer's workload for the surveillance. We conducted experiments using 8.5 hours surgery video data to confirm effectiveness of the proposed system. The system could successfully reduce the data size by 68.3%.

1. はじめに

手術をカメラで映像記録することは、普段から多くの病院で行われている。これは術野のみを拡大して記録するものであり、主に手術手技の記録を目的としている。これに対して南部ら¹⁾は、術野だけではなく手術室全体を記録し、作業内容や人の動きに潜んでいるリスクの発見にその映像を利用する提案をしている。リスクを発見することができれば、これに起因するトラブルを想定し、未然に防止策を講じることが可能となり、手術は今よりもさらに安全なものになる。

しかし、この取り組みを普及させるためには2つの大きな課題がある。1つは、リスクの発見にかかる手間である。手術が長時間に及ぶ場合、手術の映像を最初から最後まで高い集中力を維持しながら目視で分析するのは非常に手間がかかり、現実的とはいえない。たとえば、脳腫瘍手術では手術時間が10時間以上に及ぶものもある。また、作業内容や人の動きに潜んでいるリスクを発見するには、複数のカメラで多面的に手術室の状況を記録する必要がある。この場合、手術映像は手術時間×カメラ台数分となり、さらに分析の手間がかかる。

もう1つの課題は、手術映像のデータ量増加である。手術の状況を多面的に記録した映像は、手術室という閉ざされた空間の情報を第三者に提示する重要な素材にもなる。そのため、すべての映像を長期間保存することが求められるが、これら映像のデータ量は膨大である。病院としては、記憶媒体の増設コストは極力抑えたいため、映像を保存する際に符号化効率を高く設定して、記憶媒体を節約する方法が考えられる。しかしこれは、映像を詳細に分析する観点から考えると、映像が鮮明に記録されていないことが分析上の難点になりかねない。

そこで本稿では、上記2つの課題を解決する手術映像保存システムを提案する。提案システムでは手術映像中の出現頻度が低い動作にリスクが潜んでいると仮定し、我々が提案したシーン適応型動画圧縮法²⁾を適用することで、上記の課題を同時に解決する。具体的に提案システムでは、リスクの発見につながるシーン(出現頻度が低い動作シーン)ほど高品質に保存し、その他のシーン(出現頻度が高い動作シーン)を低品質に保存することで、

^{†1} 筑波大学大学院
University of Tsukuba

^{†2} 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST)

手術映像全体のデータ量を削減する．それに加え，リスクの発見につながるシーンに「しおり」をつけることで，分析の手間を大幅に低減させる．これにより提案システムは，多面的に手術室を記録した映像をリスクの発見に利用する取り組みの普及に貢献できる．

以下，2章に手術映像保存システムについて述べ，3章ではシーン適応型動画圧縮法について説明する．4章では開発した提案システムの概要について説明した後，5章でシーン分析精度の評価とデータ量削減効果の評価を行うための実験とその結果を述べ，6章で実験結果に対する考察をし，最後に7章でまとめる．

2. 手術映像保存システム

2.1 既存の手術映像保存システム

多くの病院では，カメラを用いて術野を映像記録している．この映像は，術式の共有や若手医師の育成などに利用されるほかに，患者やその家族への説明資料に利用されている³⁾．また最近では，医療事故に対する訴訟対策にも利用するケースが増えていることから，術野映像を保存することの重要性は増している．こういった状況のなか，術野映像を保存するためのシステムが製品化されているが，これらは主に映像の閲覧や検索の利便性に着目したもののばかりである．これに対して，多面的に手術室全体を記録した映像をリスクの発見に利用する提案がなされるなか，事後分析の利便性に着目した手術映像保存システムは開発されていない．事後分析に特化したシステムを開発することにより，手術の安全性を向上させる手術映像の新たな利用方法を普及させることができる．

2.2 提案システム

本稿の提案システムで処理対象とする手術映像は，図1のように複数のカメラで多面的に記録されたものである．これらの映像は，集中力の障害要因にならぬよう，医者らの視界に入りにくい天井や壁の上部に設置されたカメラで記録されている．従来の手術映像のような術野を拡大した映像には手術台周辺の情報が含まれておらず，この映像を作業内容や人の動きに含まれているリスクの発見に利用することは不可能であった．これに対して，図1のような手術映像には手術室でいったい何が起きたのかを把握できる情報が含まれているため，これらの映像を作業内容や人の動きに潜んでいるリスクの発見に利用することが可能となった．

ここで手術室におけるリスクは，小さなトラブルや手術工程の区切りなどのイベントに潜んでいるとされている⁴⁾．小さなトラブルとは，たとえば何かが台から落ちる動作や人が急いで動き出す動作などであり，これらの動作は出現頻度が低い．また，手術工程の区切りなどのイベントには手術機器の移動が含まれており，これは手術全体を通して出現頻度が低

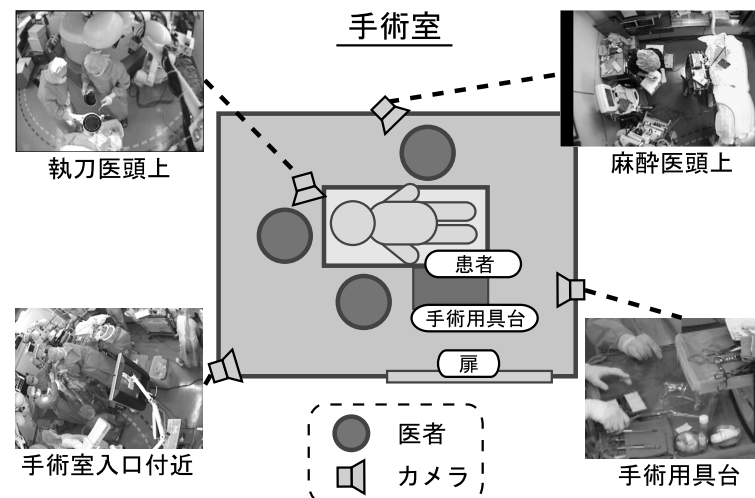


図1 複数のカメラによる手術の記録例

Fig. 1 Example of surgery video recorded with multiple cameras.

い．そこで提案システムでは，手術映像中の出現頻度が低い動作にリスクが潜んでいると仮定する．提案システムでは，手術映像に含まれるこれらのトラブルやイベントなどのリスクの発見に重要なシーンに「しおり」を付けて高品質に保存して提示し，その他のシーンは低品質に保存することで映像全体のデータ量を削減する．手術映像を分析する際には提示された高品質なシーンのみを分析すればよいので，リスクを発見する手間が大幅に低減される．また，分析に必要なシーンが低品質に保存されデータ量が小さくなるので，必要な記憶媒体のコストも低く抑えることができる．

3. シーン適応型動画圧縮法

提案システムでは，動画圧縮技術として我々の提案したシーン適応型動画圧縮法²⁾を用いる．シーン適応型動画圧縮法は，動画に含まれるシーンの重要度を算出し，この重要度を既存圧縮手法の符号化効率に反映させて圧縮処理を行う手法である．具体的には，学習用動画からフレームごとに特徴量を抽出し，これに主成分分析を用いることで出現頻度が高い動作の特徴空間内での分布を学習する⁵⁾．次に，この分布に基づき圧縮対象動画から抽出した特徴量を評価し，その逸脱度をシーンの重要度として算出する．この算出した重

要度に応じて、動画像の各シーンに施す符号化効率を変化させる。以下にシーンの重要度について説明し、そのあと提案システムにおける分布の学習に関する改良点について述べる。

3.1 シーンの重要度の算出

シーン適応型動画像圧縮法²⁾では統計的なシーン定義に基づくシーン分析を適用し、シーンの重要度を算出する。出現頻度が高い動作シーンほど重要ではないシーン、出現頻度が低い動作シーンほど重要なシーンとしてシーンを定義することで、主観によらない客観的な分析を可能とした。これにより、主観的な思い込み起因する問題を避けることができる。たとえば、「歩く」という動作は、一般的には重要なシーンであるとは考えにくい。しかし、「走る」動作がつねに起きる環境下においては、「歩く」動作の方が重要なシーンになる可能性がある。統計的なシーン定義に基づくシーン分析では、環境に応じて適応的にシーンの意味合いが定まるため、ユーザが前もって重要なシーンを想定し、列挙する必要がなくなる。本提案システムでは、手術映像中の出現頻度が低い動作にリスクが潜んでいると仮定したことから、手術映像中でリスクが潜んでいると思われるシーンほど重要度が高く算出される。以下に、重要度の算出方法について説明する。

3.1.1 立体高次局所自己相関 (CHLAC) 特徴

重要度算出のために動画像から抽出する特徴量としては、立体高次局所自己相関 (CHLAC: Cubic Higher-order Local Auto-Correlation) 特徴^{6),7)}を用いた。CHLAC 特徴は、顔画像認識などに有効な高次局所自己相関 (HLAC) 特徴⁸⁾を拡張し、時間方向の相関も加えた特徴である。CHLAC 特徴 (251 次元のベクトル) は、物の形と動き情報を表現できる特徴であり、対象の切り出しが不要で、かつモデルを用意する必要がなく、計算量が少ないという利点を持つ。HLAC 特徴や CHLAC 特徴を用いた認識技術では高精度な認識性能が報告されており、様々な分野での応用が検討されている⁹⁾⁻¹¹⁾。

3.1.2 部分空間の構成

動画像に含まれる出現頻度が高い動作は、抽出された CHLAC 特徴ベクトルが 251 次元の特徴空間内で、ある限定された領域に集中して分布することになる。ここでは、そのような分布を部分空間により近似する⁵⁾。そのため、主成分分析により主成分ベクトル $U = [u_1, \dots, u_{251}]$, $u_i \in R^{251}$ ($i = 1, \dots, 251$) を求める。主成分ベクトル U は自己相関行列 R を用いて、

$$RU = U\Lambda \quad (1)$$

の固有値問題より求まる。固有値行列 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{251})$ に基づいて、累積寄与率 η_k ($0 \leq \eta_k \leq 1$) は、

$$\eta_k = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^{251} \lambda_i} \quad (2)$$

と表され、累積寄与率 η_k が C (たとえば $C = 0.99$) となる次元までの固有ベクトル u_1, \dots, u_k により張られる空間を、出現頻度が高い動作の部分空間として採用する。

このようにして得られた直交基底 $U_K = [u_1, \dots, u_k]$ によって張られる部分空間への射影子は $P = U_K U_K'$ として表され、それに対する直交補空間への射影子は、 I_{251} を単位行列として $P_{\perp} = I_{251} - P$ となる。このとき、入力 X の上記部分空間への垂直距離 d_{\perp} は、直交補空間への射影成分で表され、

$$d_{\perp}^2 = \|P_{\perp} X\|^2 \quad (3)$$

$$= X'(I_{251} - U_K U_K')X \quad (4)$$

と表すことができる。本稿では、この距離 d_{\perp} を出現頻度が高い動作かどうかの指標として、重要度とよぶ。この重要度が大きいほど、めったに起こらないシーンであることを意味し、リスクが潜んでいると想定される重要なシーンと判断できる。

3.2 重要度に応じた符号化効率

シーン適応型動画像圧縮法では、上述した方法で検出される重要度 d_{\perp} を用いて動画像の符号化効率を適応的に変化させてデータ量を大幅に削減する。具体的には、重要度に応じて符号化効率を変化させるために、重要度と符号化効率を反比例関係とする。符号化効率を変化させるパラメータには、解像度やフレームレートなどが考えられるが、今回は簡単化のためフレームレートのみ可変にする。具体的には図 2 のように、検出された重要度とフレームレートを比例関係とする。つまり、重要なシーンほど多くのフレームで構成され、重要ではないシーンほど少ないフレームで構成される。なお、図 2 のように検出される重要度の最大値を最大フレームレートに固定すると、ユーザの意図に反し、圧縮のかけすぎやかけなさすぎが発生する可能性がある。たとえば、突出して高い重要度が 1 カ所検出された場合や、すべて同じような低い重要度が検出された場合である。そこで、図 3 に示すように、最大フレームレートを割り当てる重要度の下限ポイントは可変となっている。これにより、ある程度の選択性をユーザに与えることが可能となる²⁾。

ここで、重要度を量的に扱い符号化効率を変化させることは、手術映像を詳細に分析する観点から考えると非常に合理的である。なぜならば、リスクの発見につながる重要なシーンの取りこぼしを防ぐことができるからである。仮に重要度を質的に、つまり重要か重要でないかの二値で扱い符号化効率を極端に変化させてしまうと、リスクの発見につながる重要

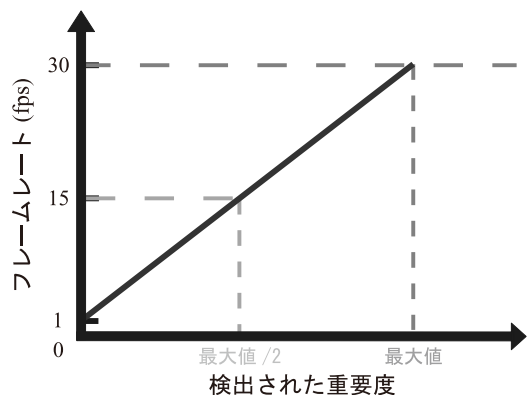


図 2 重要度に応じて割り当てられるフレームレート

Fig. 2 Assignment of frame rate corresponding to the importance value.

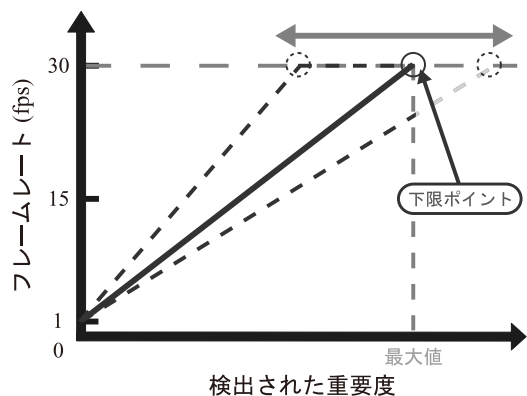


図 3 最大フレームレートを割り当てる下限ポイントの調整

Fig. 3 Adjustment of the lowest point in the importance value where the maximum frame rate is assigned.

なシーン以外のシーンすべてが低品質で保存されてしまう。出現頻度の高い動作が含まれているシーンは、リスクの発見に絶対に必要のないシーンであると断定できない。つまり、重要度を質的に扱うことは、リスクの発見を妨げる要因になりかねない。一方重要度を量的に扱うことにより、リスクの発見につながる重要なシーン以外のシーンも、段階的な品質で

保存することができる。そのため、たとえ提案システムがしおりを付与した以外のシーンを分析する際にも、極端に支障が生じることはない。

3.3 提案システムでの改良点

提案システムでは手術映像を処理対象とするにあたり、文献 2) で提案したシーン適応型動画圧縮法における部分空間の学習を、重要度算出処理を行いながら逐次更新する適応学習に改良する。手術室は手術の進行とともに状況が変わる環境であり、手術映像に含まれる動作の種類や動作の見え方が変化する。ここで、学習用動画から学習した部分空間を固定すると、手術の進行によって出現頻度が大きく変化する動作に対応することができない。そこで適応学習を用いることにより、手術室の状況の変化にも対応しながら重要度を算出できる。

具体的には、圧縮対象動画から 1 本 (1 行 251 列) の CHLAC 特徴を抽出する処理を 1 ステップとしたとき、 N ステップごとに部分空間の構成に必要な固有値問題における自己相関行列 R_{NEW} を、

$$R_{NEW} = \beta R_{OLD} + (1 - \beta)x_{(N)}x_{(N)}^T \quad (5)$$

より求める。ここで、 R_{OLD} は更新前の自己相関行列、 β ($0 < \beta < 1$) は重み係数、 $x_{(N)}$ は N ステップ間 N 本 (N 行 251 列) の CHLAC 特徴量である。この更新した R_{NEW} を用いて N ステップごとに固有値問題を解くことで、適応的に部分空間を更新する。

4. 開発した手術映像保存システム

開発したシステムは、シーン分析プロセスと、適応型圧縮処理プロセスの 2 つのプロセスからなる。図 4 にシステム概要図を示す。

4.1 シーン分析プロセス

シーン分析プロセスでは、圧縮対象手術映像から重要度を算出する。具体的には、前処理として学習用手術映像と圧縮対象手術映像をグレースケール (256 階調) に変換し、フレーム間差分法を施す。フレーム間差分法を施した手術映像には、背景の明るさの変化などによるノイズが含まれているので、二値化処理も行う。この二値化が施された学習用手術映像から CHLAC 特徴を抽出し、学習用手術映像内の頻繁に起こる行動の部分空間を学習しておく。ここで、前もって部分空間を学習しているのは、圧縮対象映像の冒頭のシーンから重要度の算出を行うためである。次に、前処理が施された圧縮対象手術映像から CHLAC 特徴を抽出し、主成分分析により構成された部分空間を用いて重要度を算出する。なお並行して、3.3 節で記したように適応学習をしながら部分空間を逐次更新する。

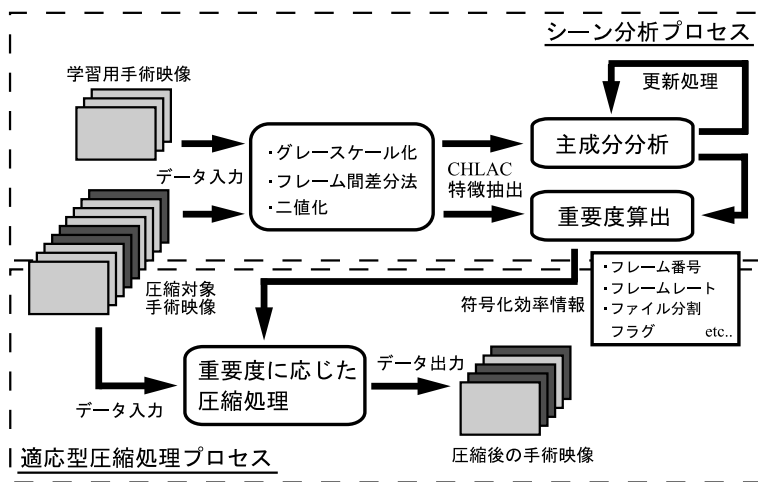


図 4 提案システムの概要
Fig. 4 Schema of the proposed system.

次に、1 秒ごとの代表重要度として 1 秒区間内の重要度の最大値を算出し、この代表重要度から図 3 の割当てにより 1 秒ごとにフレームレートを算出している。開発したシステムでは、割り当てるフレームレートの最低値を 1 fps、最大値を圧縮対象動画が元々持っている fps としている。なお、システムの仕様上、割り当てられるフレームレートの種類は圧縮対象動画のフレームレートの約数（整数）のみとなる。たとえば、圧縮対象動画の元々のフレームレートが 30 fps の場合、割り当てるフレームレートの最低値が 1 fps、最大値が 30 fps、割り当てられるフレームレートの種類は 8 種類（1, 2, 3, 5, 6, 10, 15, 30 fps）となる。

また、リスクの分析上、小さなトラブルやイベントシーンの前後も重要なシーンとなる可能性があることから、最大フレームレート（30 fps）が割り当てられたシーンの前後 2 秒にも、最大フレームレートを割り当てるようにしている。1 秒ごとに算出されたフレームレートの情報は符号化効率情報として、後の適応型圧縮処理プロセスに渡される。

4.2 適応型圧縮処理プロセス

適応型圧縮処理プロセスでは、シーン分析プロセスで作成される符号化効率情報をもとに、圧縮対象手術映像に対して変換処理を加え、既存圧縮法を施す。図 5 に適応型圧縮処理プロセスのシステム概要を示す。ここで、薄い矢印はデータが複数あることを表す。提案システムでは、分析や映像ファイルの二次利用における利便性を向上させる目的で、手術映

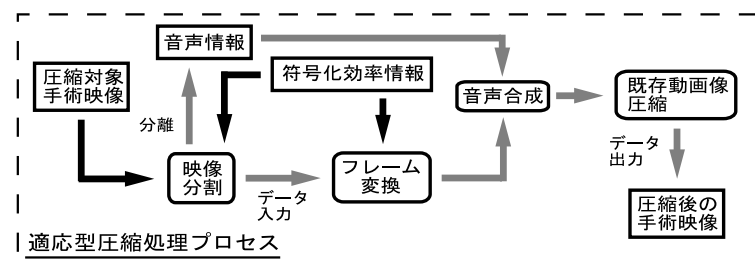


図 5 圧縮処理の概要
Fig. 5 Schema of the compression process.

像を任意の箇所で分割して保存する。そのため、圧縮処理ではまず手術映像を分割し、音声と分離する。次に符号化効率情報をもとに、指定されたフレームレートを動画像に対して適用し、データ量を削減する処理を行う。その後、分離した音声情報を結合した手術映像に既存圧縮法を施し、最終出力を行う。既存圧縮法には MPEG など一般的な圧縮方式を利用するため、提案手法を施した映像ファイル群は、一般的な動画再生ツールで一連の手術映像として再生可能となる。動画像のファイル名には、内容を見なくてもリスク発見につながる重要なシーンが含まれているファイルを把握できるよう、重要度に応じて「」印（しおり）を 5 段階で付加する。最大フレームレートを割り当てたシーンが含まれているファイルは 5 つとなる。圧縮後の映像ファイルの例を図 6 に示す。これにより、印の多いファイルから確認すればよいので、分析の手間が大幅に低減される。

5. 検証実験

実際の手術映像を用意し、提案システムの評価を行った。手術映像には図 7 のような麻酔医の頭上アングルを 2 症例分用意した。ここで、1 症例中の 1 時間分を学習用手術映像に利用し、残った 1 症例の 8.5 時間を圧縮対象手術映像として利用した。これら手術映像のフレームレートは 30 fps、フレームサイズは 360 × 240 pixel である。実験での各種パラメータは、累積寄与率の上限 C を 0.9999、適応学習の更新間隔 N を 150 ステップ、適応学習の重み係数 β を 0.85、重要度の下限ポイントを 120 とした。これらの値は予備実験の結果から定めた。また、提案システムで用いる既存動画圧縮のコーデックには MS-MPEG4 V2 を用いた。

5.1 シーン分析の検証

提案システムにおけるシーン分析の有効性を検証するために、長時間の手術映像からリス

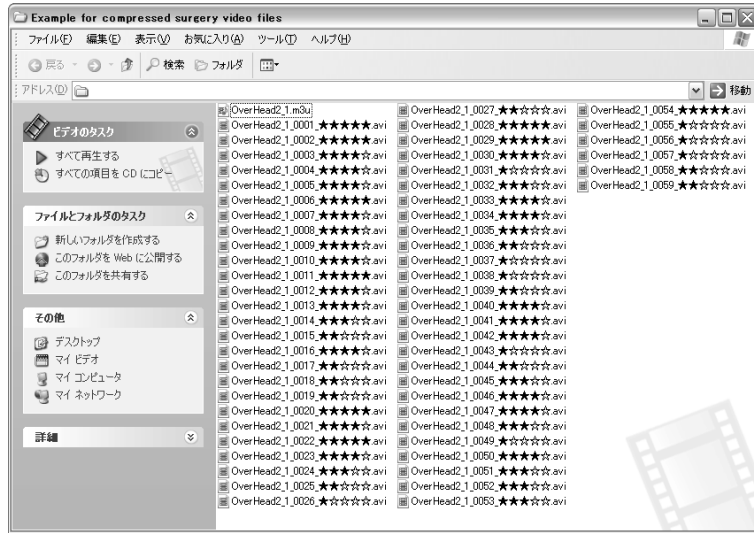


図 6 圧縮後の手術映像ファイルの例

Fig. 6 Example for compressed surgery video files.



図 7 麻酔医頭上アングル

Fig. 7 Camera angle for anesthetist.

クの発見につながるシーンをどの程度検出できているかの評価を行った．具体的には圧縮対象手術映像を目視で分析し，トラブルやイベントシーンの検出を行い，この結果と提案システムを用いて最高フレームレート（30 fps）を割り当てたシーンとの比較を行った．目視

表 1 提案システムでの検出結果（トラブル）

Table 1 Detected trouble scenes with the proposed system.

番号	出現時間	時間（秒）	検出可否	割当て秒数	動作内容
T1	0:05:24	1	○	1	床に強い光が反射する
T2	0:28:00	1	○	1	麻酔医が急に前かがみになって引出しを開ける
T3	1:45:38	2	○	2	ビニールの束が手術用具の台から落ちる
T4	1:45:57	3	○	3	麻酔医が落ちたビニールを拾う
T5	1:48:50	4	○	4	椅子に座っていた麻酔医が急に立ち上がる
T6	2:27:24	3	○	3	麻酔医が少しよろめく
T7	2:50:47	1	×	0	台の上からバインダが落ちる
T8	4:16:59	3	×	0	麻酔医が開けた箱から説明書が落ちる
T9	5:51:57	1	×	0	麻酔医が開けた箱から説明書が落ちる
T10	7:40:12	2	×	2	台の上からバインダが落ちる

表 2 提案システムでの検出結果（イベント）

Table 2 Detected event scenes with the proposed system.

番号	出現時間	時間（秒）	検出可否	割当て秒数	動作内容
E1	0:03:42	3	×	0	麻酔医が椅子を引きずる
E2	0:10:11	6	○	6	看護師がベッド上にシーツを被せる
E3	0:21:41	9	○	6	麻酔医がベッドの上にビニールを被せる
E4	1:36:32	2	○	2	手術室のライトが点く
E5	1:37:02	7	○	7	医師と麻酔医がベッドを移動させる
E6	1:37:15	2	○	2	麻酔医が手術機器 A を移動させる
E7	1:37:18	3	○	3	麻酔医が手術機器 B を移動させる
E8	1:37:24	2	○	2	麻酔医が手術機器 C を移動させる
E9	1:37:29	6	○	6	麻酔医が手術機器 C を移動させる
E10	1:37:35	1	○	1	麻酔医がベッド上の MRI 用シートを取り除く
E11	1:37:41	1	○	1	麻酔医が手術機器 C を移動させる
E12	1:37:51	1	○	1	医師がベッドの傾きを調整する
E13	1:37:52	2	○	2	麻酔医が手術機器 A を移動させる
E14	1:38:07	2	×	0	麻酔医が手術機器 A を移動させる

での分析においては，リスクの発見につながる重要なシーンとして，物が台から落ちるや，あきらかに人が何かに気づいて急な動作をとるシーンなどのほかに，手術機器の移動シーンなどを検出した．なお一連のシーンは，映像内の対象者が動作を始めるフレームから，動作が終わるフレームまでの区間とした．表 1 にトラブルシーンの検出結果，表 2 にイベントシーンの検出結果を示す．また，表 3 には提案システムが過検出したシーンの結果を示す．ここで，表における出現時刻は圧縮対象手術映像内での対象シーンの出現時刻（映像開始

表 3 提案システムでの検出結果 (過検出)
Table 3 Over-detected scenes with the proposed system.

番号	出現時間	割当て秒数	動作内容
M1	0:00:14	6	麻酔医が振り向きざまに大きく踏み込みながら歩き出す
M2	0:00:23	8	麻酔医が歩きながら手に持っている手術用具を丸める
M3	0:01:04	6	麻酔医が前かがみになって床のチューブを調整する
M4	0:01:33	5	麻酔医がベッドと壁の隙間を素早くすり抜ける
M5	0:19:21	5	麻酔医が手術機器 A の裏に手を伸ばして確認する
M6	0:19:25	7	麻酔医が手を伸ばして点滴の袋を頭上に掲げる
M7	0:53:17	5	片足重心で上体を倒していた麻酔医が上体を元に戻す
M8	1:12:25	6	麻酔医が台の上からバイндаをとる
M9	1:13:01	5	麻酔医が手に持っていたバイндаを台の上に置く
M10	1:18:09	6	麻酔医が手に持っていたバイндаを台の上に置く
M11	1:23:05	6	麻酔医が台の上からバイндаをとる
M12	1:34:17	6	麻酔医が手術機器 A にパネルを付ける
M13	1:37:17	6	麻酔医が手術機器 A と手術機器 C との狭い隙間をすり抜ける
M14	1:37:37	7	麻酔医が手術機器 C の側面部分を広げる
M15	1:37:48	7	麻酔医が手術機器 A のパネルを取り外し持ち上げて移動する
M16	1:49:00	7	麻酔医が看護師から物を受け取る
M17	1:49:17	4	麻酔医が大きく一歩踏み込んで移動する
M18	1:49:19	4	麻酔医が片足を基点に大きく一歩踏み込んで反転する
M19	2:03:48	6	手を伸ばして手術機器 A を操作していた麻酔医が椅子に座る
M20	2:04:13	6	麻酔医が手に持っていたバイндаを台の上に置く
M21	2:08:46	6	麻酔医が手に手術用具を持ったまま大きく一歩踏み込んで移動する
M22	2:36:08	6	座り込んだ麻酔医が手を伸ばして台の上から資料をとる
M23	3:07:02	6	麻酔医が手に手術用具を持ったまま大きく一歩踏み込んで移動する
M24	3:16:20	5	医師が大きく一歩踏み込んで移動する
M25	3:24:12	5	麻酔医が上体を揺らすように移動する
M26	3:24:37	6	麻酔医が大きく一歩踏み込んで移動する
M27	4:18:31	6	麻酔医が大きく一歩踏み込みながら台の上に手術用具を置く
M28	4:47:00	5	麻酔医が手にバイндаを持ちながら大きく一歩踏み込んで移動する
M29	4:54:07	5	麻酔医が椅子から立ち上がり一歩踏み込んで遠くを見渡す
M30	5:08:18	5	麻酔医が大きく一歩踏み込んで移動する
M31	5:08:22	5	麻酔医が急に振り返り大きく一歩踏み込んで移動する
M32	5:26:14	5	麻酔医が大きく一歩踏み込んで移動する
M33	6:07:28	6	麻酔医が手に資料を持ちながら一歩踏み込む
M34	6:23:19	4	麻酔医が一歩踏み込みながら手を伸ばして点滴ラベルを確認する
M35	6:23:22	5	麻酔医が手術機器 A を触ろうとしてすぐに引っ込める
M36	6:25:02	5	麻酔医が体を斜めにしながら椅子に座り込む
M37	6:32:40	6	麻酔医がバイндаを持ちながら大きく一歩踏み込んで移動する
M38	6:33:11	6	麻酔医がバイндаを持ちながら大きく一歩踏み込んで移動する



図 8 床に強い光が反射するシーン
Fig. 8 Scene of reflective intense light on floor.



図 9 ビニールの束が手術用具の上から落ちるシーン
Fig. 9 Scene of falling block of plastic on the surgical implements.



図 10 医師がベッド上にビニールを被せるシーン
Fig. 10 Scene of covering plastic on the bed by doctor.



図 11 医師と看護師がベッドを移動させるシーン
Fig. 11 Scene of moving the bed by doctor and nurse.

時を 0:00:00 とする), 時間 (秒) は対象動作を含むシーン区間の長さ, 検出可否は提案システムにおけるシーン分析の検出可否, 割当て秒数は提案システムで対象動作を含むシーン区間に最高フレームレート (30 fps) を割り当てた秒数である. 表 1 の結果から, 提案システムではトラブルシーンとして, 図 8 のような床に強い光が反射するシーン (図 8 中右端部分) や, 図 9 のようなビニールの束が手術用具の台の上から落ちるシーン (図 9 中右下部分) などが検出できた. 検出結果は全 10 シーン (21 秒) 中 6 シーン (14 秒) であり, 検出率 60% (= 6/10)・最高フレームレート割当て率約 66.7% (= 14/21) となった. ここで, 最高フレームレート割当て率とは, リスクの発見につながる重要なシーンの秒数に対する, 提案システムで最高フレームレートを割り当てた秒数である. 表 2 の結果から, 提案システムではイベントシーンとして, 図 10 のような医師がベッドの上にビニールをかけるシーンや, 図 11 のような医師と麻酔医がベッドを移動させるシーンなどが検出できた. 検

表 4 提案システムのデータ量削減効果
Table 4 Data size reduction with the proposed system.

提案システム (MB)	既存手法 (MB)	削減率 (%)
337	1,062	68.3

出結果は全 14 シーン (47 秒) 中 12 シーン (39 秒) であり, 検出率約 85.7% (= 12/14)・最高フレームレート割当て率約 83.1% (= 39/47) となった。これらの結果から, トラブルシーンとイベントシーンを合わせた最高フレームレート割当て率は約 77.9% となった。

また, 表 3 の結果から提案システムにおける過検出件数は全 38 シーン (215 秒) であった。圧縮対象手術映像は 8.5 時間 (30,600 秒) であり, 誤って最高フレームレート (30 fps) を割り当てた率は約 0.7% (= 215/30,600) と非常に小さい値となった。

これらをまとめると, 提案システムは圧縮対象手術映像 (8.5 時間) から全 56 シーン 4 分 28 秒 (トラブル 6 件・イベント 12 件・過検出シーン 38 件) をリスクの発見に最も重要なシーンとして提示した。これにより, 提案システムは手術映像を最初から最後まで目視で分析する場合にかかる多大な手間を, 大幅に低減することが可能であると確認した。これらの結果の詳細については 6 章で考察する。

5.2 データ量削減効果の検証

提案システムの有効性を検証するため, どの程度データ量が削減されるかの実験を行った。前述のシーン分析の結果をもとに手術映像を圧縮した結果を表 4 に示す。表中の既存手法とは, 提案システムを施さずに, 提案システムで用いた既存圧縮法 (MS-MPEG4 V2) のみを施した手術映像のことである。表 4 から提案システムは既存圧縮法を施した手術映像よりさらに, 約 68.3% のデータ量削減効果が得られた。この結果, 既存圧縮法を施したカメラ 1 台 (8.5 時間) 分の映像を保存するデータ容量に, 提案システムはカメラ 3 台分の映像を保存できることも分かる。

6. 考 察

5.1 節のシーン分析の検証において, トラブルシーンでは計 4 シーン (T7, T8, T9, T10) が検出できなかった。まず T7, T10 (台の上からバインダが落ちるシーン) が検出できなかった原因については, 類似したシーンの出現頻度が高いことが考えられる。手術において患者のバイタルサインなどを手術経過とともに記録することはあたりまえに行われており, 圧縮対象手術映像にも麻酔医がバインダ上の紙に記録をとっているシーンが幾度となく見受けられた。これにともない, 麻酔医が何度もバインダを移動させており, これらの動作から

バインダの動きの特徴が部分空間に学習されたものと考えられる。改善としては, 本システムでは出現頻度の高い動きを学習する部分空間を画面全体で 1 つだけ用いたが, これを複数個にする方法が考えられる。たとえば, 手術映像内でバインダが保管されている台の領域だけで部分空間を構成し, 並列して検出を行えばよい。こうすることで, 台の領域ではバインダの落下は頻度が低い動作となり, 重要なシーンとして検出されることが期待される。

次に T8, T9 (麻酔医が開けた箱から説明書が落ちるシーン) が検出できなかった原因については, 検出対象の物体が小さすぎたことが考えられる。麻酔医が開けた箱は両手サイズの小さい箱であり, それに含まれていた説明書はさらに小さいものであった。このような小さな物体からシーンの検出に有益な特徴をとることは非常に困難であり, カメラの増設やアングルを工夫するなど物理的な対策が必要と考えられる。

イベントシーンでは計 2 シーン (E1, E14) が検出できなかった。E1 (麻酔医が椅子を引きずるシーン) が検出できなかった原因については, すでに類似した動きの特徴が学習されていたことが考えられる。該当シーンは, 圧縮対象手術映像の冒頭部分に収録されていたシーンであり, このシーンの検出に用いた部分空間は学習用手術映像の特徴を大きく反映したものである。学習用手術映像には椅子の動きの特徴をとれるシーンが多く含まれていたため, 圧縮対象手術映像の冒頭では検出できなかったと考えられる。

E14 (手術機器 A を移動させるシーン) が検出できなかった原因については, 出現頻度が高くなり適応学習により手術機器 A の特徴が部分空間に反映されたことが考えられる。手術機器 A は麻酔薬を一定量注入する役割を担っている機器であり, 圧縮対象手術映像に手術機器 A の位置を調整するシーンが何度も出現するため, 部分空間に手術機器 A の特徴が反映されたものと考えられる。E1, E14 が検出できなかった原因は出現頻度を基準に重要なシーンの検出を行う提案システムの性質であり, 手術機器の移動で頻度の高いものを検出する必要が実際にあるのかどうか調査する必要がある。

過検出したシーンについて表 3 をみると, いずれも手術映像全体を通して出現頻度が低い動作であったため検出されたと考えられる。全手術映像時間 (8.5 時間) を考えると, この程度の過検出 (全 38 シーン 215 秒) は問題ないレベルと判断できる。

検出した結果について, 映像提供元の東京女子医科大学の脳神経外科医および工学系研究者とディスカッションを行った。トラブルシーンについては, T3, T4 が術中のインシデントを的確に検出しており, また, イベントシーンについては E2~E13 すべてで, 手術工程の区切りを認識する目安となる事象が検出されているとの結果が得られた。

7. おわりに

本稿では、複数のカメラで手術室の状況を多面的に記録した映像をリスクの発見に利用するうえで課題であった、長時間映像の分析にかかる多大な手間と保存すべき映像のデータ量増加を同時に解決する手術映像保存システムを提案した。具体的に提案システムは、出現頻度が低い動作シーンほど高品質に、出現頻度が高い動作シーンほど低品質に保存することで、手術映像全体のデータ量を削減した。さらに高品質に保存された出現頻度が低いシーンにはリスクが潜んでいると仮定し「しおり」を付与することで、事後チェックの手間を大幅に低減した。

提案システムを用いた実験の結果、8.5 時間もの手術映像に対して、全体のデータ量を約 68.3%も削減させることに成功した。また提案システムはリスクの発見に最も重要なシーンとして、全 8.5 時間中でわずか 4 分 28 秒、計 56 シーンにしおりを付与した。これにより、手術映像の分析にかかる多大な手間を大幅に低減できることを確認した。また実験では、しおり付与の精度についても評価を行った。その結果、目視の結果検出すべきとしたトラブル・イベントシーンの約 77.9%の秒数に最高フレームレート (30 fps) を割り当てることができた。残りの約 22%についても原因を考察し、今後改善できる見通しを得た。また、誤って最高フレームレートを割り当てた秒数は手術映像全体 (8.5 時間) の約 0.7%と、非常に小さい値であり実用上は問題ないレベルである。これらの結果より、手術映像中の出現頻度が低い動作にリスクが潜んでいると仮定した提案システムの有効性が実証された。

今後の課題として、今回実験に利用した手術映像 (麻酔医頭上) とは違った場所を記録した映像 (執刀医頭上など) を用いて実験し、提案システムの有効性を評価する。多面的に手術室の状況を記録した映像はどれも頭上アングルではあるが、記録されている動作の種類が変わることが予想され、このことが提案システムの検出率にどのように影響するかを評価する必要がある。またこれら評価の結果から、手術室における最適なカメラの設置位置を提案する。

今後の展望として、リスクが潜んでいる出現頻度が高い動作も検出できるシステムが必要であると考えられる。本稿の提案システムは、頻度が低い動作にリスクが潜んでいると仮定しており、リスクが潜んでいる出現頻度が高い動作は検出することができない。ここで、動作の出現頻度が高いということは、医師による外的規準を用いることが容易になる。今後は判別分析などの手法を用いて外的規準によって定められた動作を検出可能なシステムを開発し、本稿の提案システムに統合する。

また、手術中のリスクの発見に、手術室に存在する様々な情報を統合したシステムが必要であると考えられる。今回はカメラで取得した映像情報をもとに、作業内容や人の動きに含まれているリスクの発見に重要なシーンの検出を行った。今後は、映像情報だけでなく生体情報¹²⁾ や手術室内の人の位置情報¹³⁾ など、複数の情報を統合してリスクの発見に重要なシーンの検出を行うことで、より高性能なシステムの構築が期待される。

謝辞 手術室内映像は東京女子医科大学において同意のもとに記録された資料を匿名化した状態で提供いただいた。本稿の研究を遂行するにあたり、東京女子医科大学先端生命医学研究所の先生方に多くの有益なご意見をいただいた。心より感謝申し上げます。本研究は NEDO 内視鏡下手術支援システムの研究開発事業 (P10003) により実施された。

参 考 文 献

- 1) 南部恭二郎, 伊関 洋: 手術戦略デスクと手術安全支援システム, 日本エム・イー学会誌, Vol.44, No.2, pp.257-264 (2006).
- 2) 坂部史生, 村川正宏, 小林 匠, 樋口哲也, 大津展之: シーンの重要度に応じて符号化効率を変化させる適応型動画画像圧縮法, 電気学会論文誌 C 編, Vol.130, No.7, pp.1177-1185 (2010).
- 3) 松岡圭介: 医用映像支援システムのこれから, 映像情報 Industrial, No.1, pp.53-56 (2010).
- 4) 南部恭二郎, 伊関 洋: 手術安全のためのプロセス改善システム, 医科器械学, Vol.75, No.5, pp.282-287 (2005).
- 5) Nanri, T. and Otsu, N.: Unsupervised Abnormality Detection in Video Surveillance, *IAPR Conference on Machine Vision Applications*, pp.574-577 (2005).
- 6) Kobayashi, T. and Otsu, N.: Action and Simultaneous Multiple-Person Identification Using Higher-order Local Autocorrelation and Factor Analysis, *Proc. International Conference on Pattern Recognition*, pp.741-744 (2004).
- 7) Kobayashi, T. and Otsu, N.: Three-way Auto-correlation Approach to Motion Recognition, *Pattern Recognition Letters*, Vol.30, pp.212-221 (2009).
- 8) Kurita, T., Otsu, N. and Sato, T.: A Face Recognition Method Using Higher Order Local Autocorrelation and Multivariate Analysis, *IAPR International Conference on Pattern Recognition*, pp.213-216 (1992).
- 9) Mimura, S., Ito, K., Kobayashi, T., Takigawa, T., Tajima, A., Sawamura, A. and Otsu, N.: The Cow Gait Recognition Using CHLAC, *ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security*, pp.56-57 (2008).
- 10) Sakabe, F., Murakawa, M., Kobayashi, T., Higuchi, T. and Otsu, N.: Anomalously Detection for Surgery Videos Using CHLAC Feature, *ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security*, pp.66-68 (2009).

- 11) Nosato, H., Sakanashi, H., Murakawa, M., Higuchi, T., Otsu, N., Terai, K., Hiruta, N. and Kameda, N.: Histopathological Diagnostic Support Technology using Higher-order Local Autocorrelation Features, *ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security*, pp.61–65 (2009).
- 12) 荒木英人, 村川正宏, 小林 匠, 樋口哲也, 久保田一, 大津展之: 高次局所自己相関特徴による多チャンネル時系列データからの異常検知, *電気学会論文誌 C 編*, Vol.129, No.7, pp.1305–1310 (2009).
- 13) Nara, A., Izumi, K., Iseki, H., Suzuki, T., Nambu, K. and Sakurai, Y.: Surgical Workflow Analysis based on Staff's Trajectory Patterns, *The 1st Workshop Modeling and Monitoring of Computer Assisted Interventions* (2009).

(平成 23 年 2 月 7 日受付)

(平成 23 年 3 月 29 日再受付)

(平成 23 年 4 月 19 日採録)



坂部 史生

2009 年武蔵工業大学大学院 (現, 東京都大学大学院) 工学研究科博士前期課程修了。現在, 筑波大学大学院システム情報工学研究科博士後期課程在籍。画像処理の研究に従事。



村川 正宏

1999 年東京大学大学院工学系研究科博士課程修了。博士 (工学)。同年電子技術総合研究所 (現, 産業技術総合研究所) 入所。現在, 情報技術研究部門主任研究員。筑波大学連携大学院准教授, 東邦大学連携大学院准教授兼任。最適化アルゴリズム, 適応型ハードウェア, 適応学習手法の研究に従事。第 2 回進化システム国際会議最優秀論文賞, 平成 12 年度つくば奨励賞, 電気学会ミレニアム最優秀論文賞, 平成 19 年電気学会「産学連携による論文」最優秀論文賞受賞。



小林 匠

2003 年東京大学工学部機械情報工学科卒業。2005 年同大学院修士課程修了。同年 (株) 東芝入社, マルチメディアラボラトリー研究員を経て 2006 年産業技術総合研究所入所。2009 年筑波大学大学院博士課程修了。博士 (工学)。現在, 情報技術研究部門研究員。パターン認識・多変量解析の理論と応用の研究に従事。平成 19 年度 PRMU 研究奨励賞受賞。



樋口 哲也

1982 年慶應義塾大学大学院工学研究科博士課程修了。工学博士。1983 年電子技術総合研究所入所。1990~1991 年カーネギーメロン大学客員研究員。現在, 産業技術総合研究所情報技術部門主幹研究員。筑波大学連携大学院教授兼任。進化型ハードウェア, 遺伝的アルゴリズムの研究に従事。第 25 回市村学術賞, 電気学会ミレニアム最優秀論文賞, 平成 19 年電気学会「産学連携による論文」最優秀論文賞受賞。



大津 展之

1969 年東京大学工学部計数工学科卒業。1971 年同大学院修士課程修了。同年電子技術総合研究所入所。1985 年情報数理研究室室長。1991 年知能情報部長。1992~2010 年筑波大学連携大学院教授兼任。2001 年~現在, 産業技術総合研究所フェロー。同年~2007 年東京大学大学院情報理工学系研究科教授兼任。工学博士。パターン認識の基礎と応用に関する研究に従事。電子通信学会学術奨励賞, 行動計量学会優秀賞, 映像情報メディア学会丹羽高柳賞業績賞等受賞。