

[ポスター講演]映像照合による腹腔鏡手術映像へのタグ付けに  
関する検討佐藤 健司<sup>†</sup> 道満 恵介<sup>†</sup> 目加田 慶人<sup>†</sup> 三澤 一成<sup>††</sup> 森 健策<sup>†††</sup><sup>†</sup>中京大学情報科学研究科 〒470-0393 愛知県豊田市貝津町床立 101<sup>††</sup>愛知県がんセンター 〒464-8681 愛知県名古屋市千種区鹿子殿 1 番 1 号<sup>†††</sup>名古屋大学情報連携統括本部 〒464-8601 愛知県名古屋市千種区不老町E-mail: <sup>†</sup>{h11213m@st, kdoman@sist, y-mekada@sist}.chukyo-u.ac.jp, <sup>††</sup>misawakzn@aichi-cc.jp,<sup>†††</sup>kensaku@is.nagoya-u.ac.jp

あらまし 腹腔鏡下手術は低侵襲手術とも呼ばれ、患者への負担が少ない術式である。しかし、カメラから得られる視野は狭く、少ない情報を頼りに鉗子を操作するため、術者に高度な技術を要求する術式である。手術映像に対して、病巣部分の切除や縫合などの動作が起こる箇所にタグ付けすることで、手術結果の評価などに利用可能である。本稿では、腹腔鏡手術映像から特定の動作シーンを検索する手法を提案する。提案手法は、固有空間法を利用した特徴ベクトルの次元削減と、連続 DP マッチングによる類似映像検索によって構成される。腹腔鏡手術映像を用いた実験の結果、特定動作シーンの検索が可能であることを示す。

キーワード 映像のタグ付け、固有空間法、連続 DP マッチング、腹腔鏡手術

## Automatic Annotation of Laparoscopic Surgery by Video Comparison

Takeshi SATO<sup>†</sup> Keisuke DOMAN<sup>†</sup> Yoshito MEKADA<sup>†</sup> Kazunari MISAWA<sup>††</sup>  
and Kensaku MORI<sup>†††</sup><sup>†</sup>Graduate School of Computer and Cognitive Sciences, Chukyo University 101 Tokodachi, Kaizu-cho, Toyota-shi,  
Aichi, 470-0393 Japan<sup>††</sup>Aichi Cancer Center 1-1 Kanokoden, Chikusa-ku, Nagoya-shi, Aichi, 464-8681 Japan<sup>†††</sup>Information and Communications Headquarters, Nagoya University Furo-cho, Chikusa-ku, Nagoya-shi, Aichi,  
464-8601 JapanE-mail: <sup>†</sup>{h11213m@st, kdoman@sist, y-mekada@sist}.chukyo-u.ac.jp, <sup>††</sup>misawakzn@aichi-cc.jp,<sup>†††</sup>kensaku@is.nagoya-u.ac.jp

**Abstract** Laparoscopic surgery or minimally invasive surgery is few burdens to a patient. However, because of limited viewing field and difficulty of operating laparoscopic instruments, it requires a high technique to surgeons. It is effective to evaluate an operation by adding tags, which corresponds to the specific intraoperative scene like removing lesion or suturing scene. In this paper, we propose a method of specific operative scene retrieval from laparoscopic surgery videos. The proposed method employs the eigenspace method for dimensionality reduction, and continuous DP matching for searching near similar image sequences. Experimental results showed that the proposed method can retrieve specific operative scene for video tagging.

**Keyword** Video tagging, Eigenspace method, Continuous DP matching, Laparoscopic surgery

## 1. はじめに

腹腔鏡手術は、腹腔鏡と呼ばれるカメラと鉗子と呼ばれる器具を二酸化炭素にて気腹した腹腔内の空洞に挿入し、腹腔鏡で撮影される映像を手がかりに治療を行う術式である。患者の腹部に大きさが 1 から 5cm 程度のポートを空けて術具を挿入して行われるため、術後の傷は小さく、患者の社会復帰も早い。しかしな

がら、狭い視野と作業スペースで、カメラの映像だけを頼りに、複数の医師が協調しながら鉗子の複雑な操作を行わなければならないため、術者に高度な技術を要求する。また、体内に挿入されたカメラが汚れた場合には、一端体外に出して洗浄しなければならないなど、開腹手術にはない作業がある。数時間にも及ぶ手術中に撮影されるカメラ映像は、手術記録として保存

され、映像は術後の評価などに用いられる。術後の評価において検討すべきことは、患部の摘出や鉗子操作自体だけでなく、出血の回数、カメラが洗浄された回数など多岐にわたる。この評価のためには、手術映像を全て確認し、病変部摘出や出血など、特定の動作シーンに対するタグ付けを行う必要があるが、人手によるタグ付けには多大な労力がかかる。

一般の映像に対するタグ付けの研究としては、スポーツ映像、ニュース映像、調理動作映像などを対象として、幅広く研究されている[1-3]。野球やアメリカンフットボールなどのスポーツを対象とした場合には、そのスポーツ特有のルールや、映像作成時の典型的なカメラ割りや歓声の大きさなどを手がかりに解析する方法がとられる。ニュース映像を対象とする場合には、ニュースの内容を書き下したクローズドキャプションを利用することが可能であり、ニューストピックに関する時系列的な推移などが解析される。

一方、手術映像を対象とした場合、対象の部位や疾病は患者毎に異なり、様々な経緯を経て患部が摘出される事となる。例えば、同じ疾病を対象とした手術であっても患者の体型や患部の部位が異なれば、ポート位置などの条件が異なる。カメラワークなどの明確な映像文法がないため、上述の例の様な文法は適用できない。

そこで本研究では、直接的に類似映像を照合することにより、手術映像から特定の動作シーンと類似する部分を検索し、タグ付けする手法の提案と、検索可能なシーンに関する検討を行う。

## 2. 提案手法

映像照合により、手術映像から特定の動作シーン(以後、クエリと呼ぶ)と類似する部分を検索することを目的とする。手術映像とクエリの各フレームから特徴ベクトルを算出し、手術映像の各フレームとクエリの各フレーム間の距離をパラメータとして、クエリと類

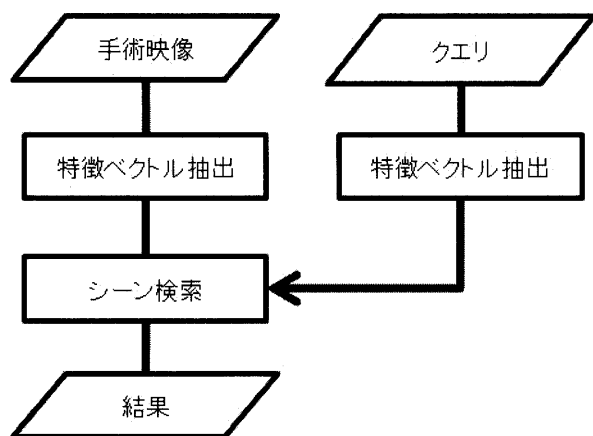


図1 提案手法

似するシーンを手術映像から検索する(図1)。

### 2.1. 特徴ベクトル算出

まず、映像の各フレームから特徴ベクトルを算出する。茅野らは、手術映像の各フレームに対して、そのフレームに写っているものを分類することを目的とし、各フレームを10種類(体外、トロカール内、肝臓、鉗子が映っているなど)に分類する研究を提案した。この研究では、分類特徴に色情報と輝度情報を使用し、部分空間法を用い手術映像の各フレームを分類した。そして、色情報と輝度情報、2種類の分類特徴においてこれらを分類可能であるとの結果を得ている[4]。本研究では特徴ベクトルとして、色情報(アピランス)を用い、固有空間法により圧縮されたベクトルを各フレームの特徴ベクトルとした。また、映像間の照合は、特徴ベクトル同士の連続DPマッチングにより実現した。以下の詳細な説明をする。

#### 2.1.1. 特徴量

分類に用いる特徴量は、(a)色情報と、(b)色情報に鉗子の有無を含めた特徴の2種類とする[5]。

鉗子は腹腔鏡手術において重要な器具であり、術者の手の代わりとして病変部の摘出や縫合を行う。したがって、鉗子は手術映像解析に重要と考えられるため鉗子の有無を特徴に加える。

鉗子は金属で構成されており、映像中では輝度飽和している割合が高い。しかしながら、体内に配置されたガーゼなども輝度飽和した画素が多い物体である。そこで、各RGB値が、

$$119 \leq R \leq 151$$

$$119 \leq G \leq 151$$

$$119 \leq B \leq 151$$

となる中輝度の画素を鉗子領域とし、この画素数を特徴ベクトルに加えた(図2(右))。

#### 2.1.2. 固有空間法による特徴抽出

映像の各フレームを分類するために固有空間法[6]を用いる。茅野らの文献[4]では10個のカテゴリに対応する10個の部分空間を生成し、未知の手術映像の各フレームを10個の部分空間に射影し、類似度(射影したベクトルのノルム)が最大となるカテゴリへ分類した。本研究では手術映像全体を使った固有空間を生成



図2 鉗子領域 原画像(左)、色成分に対する2値化による鉗子領域の抽出結果(右)

し、未知映像の各フレーム系列を固有空間内での軌跡として扱う事とした。手術映像の持つ分布を低次元で表すことができる固有空間に各データを射影する。映像をその分布の特徴を保ったまま低次元のベクトル系列と扱う事が出来るため、高速かつ適切な検索に寄与すると考える。

具体的な特徴ベクトル算出は学習と認識の2段階からなる。学習段階はまず、学習映像の各フレーム映像のサイズを縮小する。次に画素の正規化を行う。入力フレームの画素値を  $R, G, B$ , 正規化後のフレームの画素値を  $r, g, b$ ,  $S = R + G + B$  としたとき、

$$r = \frac{R}{S}, g = \frac{G}{S}, b = \frac{B}{S}$$

により、正規化した画素値を求める。

正規化された画素  $x$  の各成分を  $x_r, x_g, x_b$  とし、これらを画素数分並べたベクトルに対して、各成分の平均0, ノルム1とした正規化ベクトル  $f$  を求める。

行列  $M = [f_1^T, f_2^T, \dots, f_n^T]$  ( $n$  は画像数) から自己相関行列  $Q = MM^T$  を求め、 $Q$  を固有値展開することにより固有ベクトルを得る。そして固有値を大きい順に選択し、それらに対応する固有ベクトルにより張られる固有空間を得る。

認識段階は、学習段階と同じ方法で手術映像のフレームから正規化ベクトル  $g$  を求め、 $Mg^T$  により固有空間へフレームを射影する。この固有空間上へ射影したベクトルを特徴ベクトルとする。

### 2.2. 類似シーン検索

入力映像からクエリと類似するシーンを検索する。

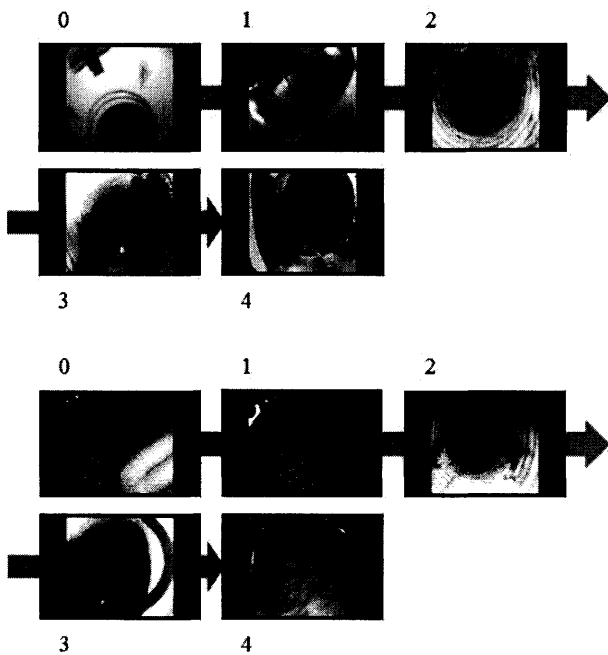


図3 クエリ(a) (上), 検索結果(下)

シーン検索には連続 DP マッチングを用いる[7]。

連続 DP マッチングは、2つの系列の時間伸縮が可能なマッチング手法であり、音声認識におけるワードスポッティングなどに利用されている。

連続 DP マッチングにおける各特徴ベクトル間の距離としては、手術映像の各フレームとクエリの各フレームが持つ特徴ベクトル間のユークリッド距離を使用する。手術映像を時間的に分割し、それらからクエリ映像を検索し、コストが小さい対応付け結果10個を検索結果とした。

### 3. 実験

実験試料は胃部分切除腹腔鏡手術3症例を用いた。そのうち、固有空間生成のために1症例、被検索対象映像(入力映像)に1症例、検索クエリは残りの1例の一部を用いた。クエリは(a)カメラが体内へ入るシーン、(b)クエリが体外へ出るシーン、(c)出血したシーンと(d)胃から脂肪を剥離するシーンの4種類である。

入力映像の長さは約60[分]~100[分]、クエリの長さは20[秒]とした。分類における固有空間の学習画像は映像から500フレーム抽出し、固有空間次元数は実験的に35次元とした。連続 DP マッチングによる検索の範囲は入力映像の3分間の区間毎とし、各区間は前後の区間と1分間のオーバーラップを持たせた。類似シーン検索においては、各被検索映像から相違度が低いものから順に10個の部分映像を検索結果とした。

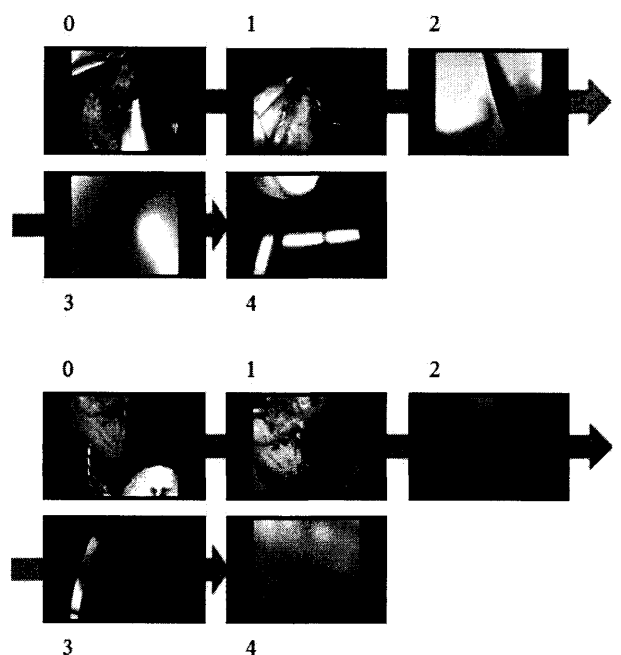


図4 クエリ(b) (上), 検索結果(下)

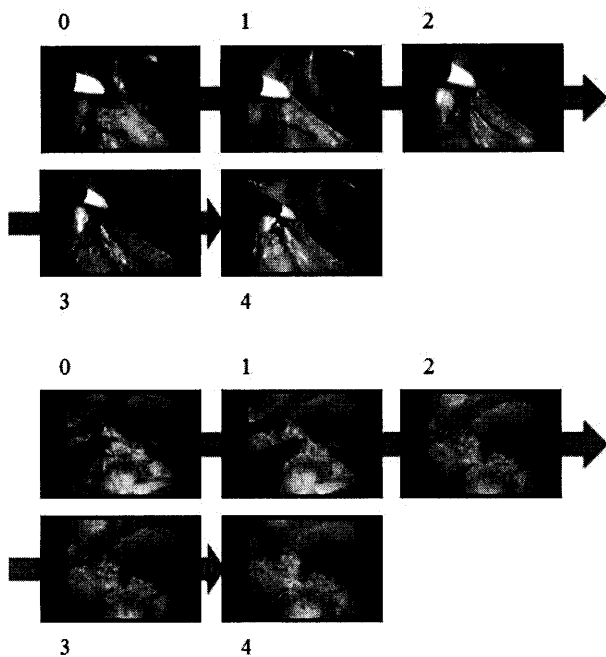


図 5 クエリ(d)(上), 検索結果(下)

表 1 各シーンに対する検索正解数

	特徴		正解箇所
	色	色+鉗子	
カメラが入る	6	5	8
カメラが出る	5	5	7
出血	2	4	4
脂肪剥離	3	6	14

4. 結果

検索結果の例を図 3 から図 5 に示す。それぞれ、カメラが体内に入るシーンのクエリと検索結果(図 3), カメラが体外に出るシーンのクエリと検索結果(図 4), 脂肪を剥離するシーンのクエリと検索結果(図 5)である。各シーンに対する検索正解数について表 1 に示す。

5. 考察

腹腔鏡カメラが体内に入ったりするシーンについては、比較的良好な検出正解率であった。これは例えばカメラが体内に入る場合では、クエリ映像の冒頭に体外の手術室の壁や患者の腹部の外観があり、次にトロカー内を通過して腹腔内に入るといふ、比較的大きな変化を含むため、固有空間内の特徴ベクトルが特有の軌跡となっていた事が予想される。また、出血のシーンでは検索対象数が少ないものの、鉗子の特徴を加える場合に全てを検索することが可能であった。ただし、相違度の小さい順に、1, 2, 3 位と 7 位が正解検索結果であった。出血のシーンとは、鉗子操作に伴い若干の血液が滲み出るようなケースであり、血管の近くで

鉗子操作をしているようなシーンを誤検出していた。脂肪の剥離については、鉗子がある程度含むつつ脂肪特有の色成分が含まれているフレーム系列が検出対象であり、正解検索結果以外にも似たような色成分が含まれている部分の誤検出が多くあった。脂肪の剥離は同じような特徴が連続するシーンであり、カメラの出入りなどの大きな映像変化がないため、連続 DP マッチングによる類似パターン検出には適していないため、今後別の手法の検討が必要となる。

6. まとめ

本稿では、映像検索により腹腔鏡手術映像から特定の動作シーンと類似するシーンを検索する手法を提案し、映像照合により手術映像からクエリと類似するシーンが検索できることを示した。フレーム内の色分布に加え、鉗子の有無に関する特徴を各フレームの特徴ベクトルに加えることで、検索精度が向上する可能性があることが確認出来た。今後は鉗子の動き特徴による誤検出削減と、大局的なシーン構造の理解を行う。

文 献

- [1] N. Nitta, N. Babaguchi, and T. Kitahashi: "Generating Semantic Descriptions of Broadcasted Sports Videos Based on Structures of Sports Games and TV Programs," *Multimedia Tools and Applications*, Vol.25, No.1, pp.59-83, 2005.
- [2] 小川晃, 高橋友和, 井手一郎, 村瀬洋, "画像とテキストの利用による同一ニュースの言語横断検索," *情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア]* 2008(27), pp.35-40, Mar.2008
- [3] 久原卓, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, "CHLAC 特徴の周期性解析による料理映像中の繰り返し調理動作区間の抽出と識別," *電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎* 110(457), pp.61-66, Feb.2011
- [4] 茅野貴洋, 目加田慶人, 森健策, 三澤一成, "部分空間法による腹腔鏡手術シーン分類の検討," *日本コンピュータ外科学会誌*, Vol.12, No.3, pp.452-453, Nov.2010.
- [5] 佐藤健司, 目加田慶人, 森健策, 三澤一成, "映像照合による腹腔鏡手術シーン検索に関する検討," *日本コンピュータ外科学会誌*, Vol.14, No.13, Nov.2012
- [6] 出口光一郎, 岡谷貴之, "固有空間法はなぜうまく働くか," *情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア]* 2001(66), pp.1-8, July.2001
- [7] 齊藤剛史, 内田克彦, 小西亮介, "連続 DP マッチングを用いた発話シーンからの単語スポットティング認識," *電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解* 110(219), pp.53-58, Oct.2010