

多視点手術動画の自動視点切替

大石 圭^{*1} 斎藤 英雄^{*1} 梶田 大樹^{*2} 高詰 佳史^{*2}

Camera Switching for Multi-view Surgery Videos

Kei Oishi^{*1}, Hideo Saito^{*1}, Hiroki Kajita^{*2} and Yoshifumi Takatsume^{*2}

Abstract — 手術のビデオ撮影による記録の有用性は以前から認識されてきた。すでに多くの手術室では固定式の術野カメラが設置されている。しかし、実際には外科医の頭部や体によって術野が隠れてしまいカメラによる手術の記録が困難である。我々はこの問題を解決するために、無影灯に複数のカメラを取り付け記録するシステムを提案した。また、術野を認識し自動的にカメラを切り替えることで、頭部の映り込みの少ない映像の生成を試みる。その際、一般的に頻繁なカメラの切り替えは、動画の視聴品質（Quality of View, QoV）が低下する。そこで我々は、組み合わせ最適化の手法で、最短経路問題で広く用いられるダイクストラ法をこの問題に応用し、一定時間カメラの切り替えが行われず、動画全体で観測している術野領域が最大になるカメラスケジューリング手法を提案する。

Keywords : カメラスケジューリング, ダイクストラ法, 多視点カメラ, 術野ビデオカメラ, ビデオ解析

1 はじめに

手術をビデオ撮影し記録することの有用性は、以前から認識されている。多くの手術室ではすでに固定式の術野カメラが設置されている。しかし実際には、カメラと術野の間に外科医の頭部が入り込んでしまうため、既存の単一の術野カメラでは良好な手術映像を記録することは困難である。

そこで我々は、外科医が手術中に无影灯の照明によって術野を照らす時に、无影灯に取り付けられている多数の照明のうち、少なくとも一つの照明が術野を照らすことに着目し、无影灯の各照明に対応する位置にカメラを取り付けることにより、少なくとも一つのカメラで術野が確実に撮影されるようにしたシステムを考案した。本稿では、このシステムにより撮影された多視点動画像に対して、術野領域の撮影範囲の大きさが出来るだけ大きくなるように自動的にカメラを切り替える（視点を切り替える）手法を提案する。本手法は、以下の2つの課題を解決するものである。

- 術野領域に注目したカメラスコアリング
- 複数カメラの切り替え

最後にマルチカメラ搭載型无影灯で料理を撮影し、その映像での実験を通して、提案手法の有用性について議論する。

2 関連研究

本論文ではカメラ切り替えによる手術の新たな記録手法を提案する。そこで節2.1では、これまでになされた手術の記録のための試み、そして節2.2では、カメラ切り替え技術を紹介し、提案手法の位置付けを明確にする。

2.1 手術の記録

医師は後世に技術を伝える義務を有する。そのためにも、手術を記録し訓練映像を生成することは重要である。また、執刀医の手術の復習、訴訟対策等の観点からも手術の記録の有用性は、以前から認識されている。腹腔鏡下手術など、元々カメラ越しに手術を行うものは、その手術の記録は容易である。しかし、切開が伴う手術など、医師が直接視認する手術の場合、外科医の存在や空間的な制約等によってその記録は困難である。

Kumar[1]らは、カメラアームをデザインし、それにカメラを搭載することで、手術の撮影を行った。カメラアームは、執刀医の邪魔にならない位置に設置され、術野から遠い位置に設置されることが多い。また、状況に応じたカメラの位置操作が煩わしい。そこでByrd[2]らは、无影灯にカメラを搭載することで、上記の問題解決を図った。しかし、医師の頭部や体の映り込みによって、視界が悪く、術野を常に観測し続けることは難しい。

医師の目の間に設置されたカメラによって手術を記録する試みも行われてきた。そのような記録システムは高解像度ではなく、ルーペ等の追加コストによって高価になることが多い[3, 4]。Akshay[5]らは、高解像度なカメラ（GoPro Hero 4）を医師の頭部に取り付け

^{*1}慶應義塾大学大学院理工学研究科

^{*2}慶應義塾大学医学部

^{*1}Graduate School of Science and Technology, Keio University

^{*2}Faculty of Medicine Department, Keio University

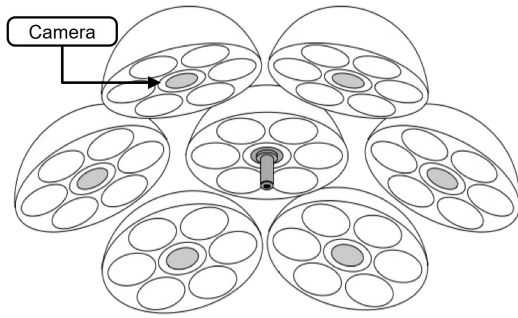


図1 マルチカメラ搭載型无影灯のイメージ
Fig.1 Image of multi-camera mounted shadowless lamp

ることにより、手術の記録を行った。医師の頭は、手術中大きく動くものであり、しばしばその映像は視聴者にとって気分を害するものになり得る。また、医師の頭部に取り付けることは、医師に不快感を与え、手術の妨げになりうる。

一方、提案手法では、无影灯に備えられた複数の光源に対応する形で複数台カメラを設置し、手術を記録する。それによって、医師にカメラを意識させることなく術野を記録するとともに、カメラを自動的に切り替えることで、術野以外の映り込みを抑えた手術の記録を実現する。

2.2 カメラ切り替え

近年、オフィス環境、スポーツスタジアム、街中などあらゆる場所に複数台のカメラを取り付けることが一般的になっている。しかし、その膨大な情報の中から必要な情報だけを取り出すことは難しい。そこで、カメラの自立制御技術や自動視点切り替え映像の生成技術、ハイライト映像の生成技術は重要な課題と位置づけられている [6]。これまでに、スポーツシーン [7, 8, 9] や監視システム [10, 13] 等への応用が提案されている。自動視点切り替え技術の課題は大きく2つに分けることができる。第一に、シーンに応じたビデオ分析。第二に、目的を達成するための最適な視点の選択である。視点切り替えによって自動生成されるビデオコンテンツの品質を一般化することは難しく、その品質は対象のシーンとユーザの目的に依存する [11]。

Liu[12] らは、プロの動画編集者に動画編集のためのインタビューを行い、そのルールを実装した。そのルールに基づいて、会議映像において、発言者、聴衆、全体の3台のカメラの切り替えを行った。Doubek[13] らは、オフィス環境において固定カメラネットワークを使用して、移動する対象の観測を行った。カメラの選択は、各カメラのスコアに基づいて行われるが、そのスコアに有意な変化があった場合のみカメラの切り替えが行われるよう、抵抗係数が導入されている。

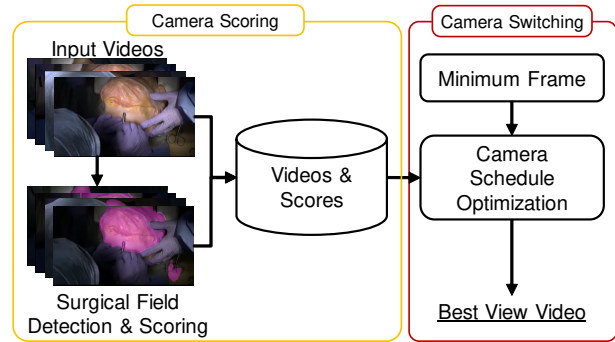


図2 提案手法の流れ
Fig.2 Flow of the proposed method

しかし、こうしたカメラの切り替えは、スコアの競合するカメラが存在した場合、カメラの切り替えが頻繁に行われ、動画の視聴品質 (Quality of View, QoV) が低下する可能性がある。

Jiang[9] らは、カメラの切り替え回数を制限するために、ターゲットの大きさ、姿勢、向き等により算出される QoV に依存したコスト関数を提案し、頻繁なカメラの切り替えを制御した。Daniyal[8] らは、物体の視認性スコアを多変量ガウス分布モデルを用いて算出し、カメラの切り替えに部分観測マルコフ決定過程 (POMDP) を用いることで、カメラの切り替えを抑えながら視認性スコアの最大化を図った。これらの処理は逐次処理であり、過去と現在の情報からカメラを選択するため、しばしば、最適な選択とは言えない場合が存在する。また、各カメラで算出される QoV と切り替えのための重みを最大化または最小化するため、ユーザの指定する最適と思われる切り替えのための重みは QoV に依存し、対象シーンによって変わると考えられる。さらに、医師のような専門家ではない場合、重みによるカメラの切り替えの制御は、出力の結果を想像し難いため、重みの決定が難しい。

一方、提案手法では、最小フレームを指定し、フレーム全体で最適化を行うため、カメラの切り替え頻度は対象シーンによって変わらない。また、医者などの専門家でない場合でも、容易にパラメータ設定ができる。

3 提案手法

外科医は術野が常に光源によって照らされている環境下で手術を行う。そのため、无影灯には複数の光源が取り付けられており、外科医の頭部や体によっていくつかの光源が遮蔽されたとしても、ほかの光源が補完し、常に術野が照らすように作られている。そこで提案手法では、図1に示すように、无影灯に備えられた光源に対応する形で複数台カメラを設置することで、術野が明るい限り、常にいずれかのカメラによって撮

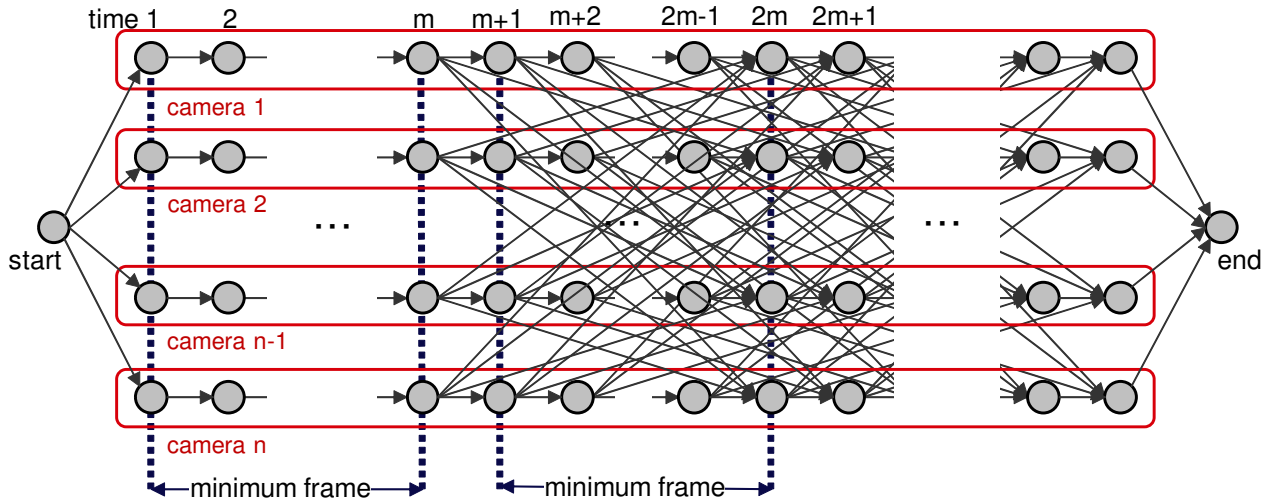


図3 最適化のためのグラフ生成
Fig.3 Generate graph for camera schedule optimization

影が完遂される状況を作り出す。このカメラシステムを用いて、術野のみを撮影しているカメラが自動的に選択される、カメラ切り替え映像を生成する。この際問題となる、カメラのスコアリング方法とスイッチング方法について、それぞれ3.1節、3.2節にて述べる。また、提案手法の流れを図2に示す。

3.1 カメラスコアリング

カメラの切り替え映像を生成するために、その基準となるカメラのスコアリングが重要となる。提案手法では術野を認識しその領域に基づいてカメラのスコアリングを行う。術野のセグメンテーションを行う検出器 D を以下のように定義する。

$$D(I(i, j)) = \begin{cases} 1 & \text{surgeical field} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

このとき I は入力画像である。

術野は時間経過とともに形が変わり、その領域を検出し続けることは難しい。そこで、Cheng[14]らの手法を用いて術野のセグメンテーションを行う。Chengらは色とテクスチャの情報から手のセグメンテーションを行った。この手法は色とテクスチャの情報を学習するため、U-NET[15]に代表されるようなディープラーニングの手法に比べ学習画像の枚数が少なくてよい。認識したい術野は体表と体内であり、それぞれ場所や形が変わったとしても、色やテクスチャに関して大きな差はなく、Chengらの手法と相性がよい。

提案手法では、体表と体内に対しそれぞれ検出器 D_1 、 D_2 を用意し、それぞれの検出器で検出された領域の論理和を術野とする。このとき、あるカメラ c の

時刻 t におけるスコア s_t^c を以下の式のように定義する。

$$s_t^c = 1 - \frac{\sum_{i=0}^{w-1} \sum_{j=0}^{h-1} D_1(I_t^c(i, j)) \cup D_2(I_t^c(i, j))}{wh} \quad (2)$$

ここで、 w は画像の幅、 h は画像の高さである。得られたスコアと動画は図2に示すように、後述のカメラスイッチングのために、データストレージに保存される。

3.2 カメラスイッチング

n 台のカメラ c に対し、ある時刻において算出されたスコアが最もよいカメラの映像を出力し続けた場合、カメラの切り替えが頻繁に行われるため、著しく動画の視聴品質（Quality of View, QoV）が低下し、ユーザは不快感を感じる。そのためカメラのスイッチングを制御する必要がある。提案手法では、ユーザが指定する最小フレーム間カメラの切り替えが行われないかつ、検出した術野の最大化、すなわち術野以外の領域の最小化を図る。そこで、組み合わせ最適化の手法で、最短経路問題で広く使われるダイクストラ法[16]を応用した手法を提案する。

3.2.1 グラフの生成

提案手法では、組み合わせ最適化問題でダイクストラ法を使うにあたり、最適化のためのグラフ生成を工夫することによって、ユーザが指定する最小フレーム間カメラの切り替えが行われないかつ、術野以外の領域の最小化を行う。期待されるグラフの生成結果を図3に示す。各ノード V はカメラ番号 c 、時刻 t の情報をもつ。その情報によってノード間でエッジをつなぐか否かが決定され、その重みも変化する。エッジ E を

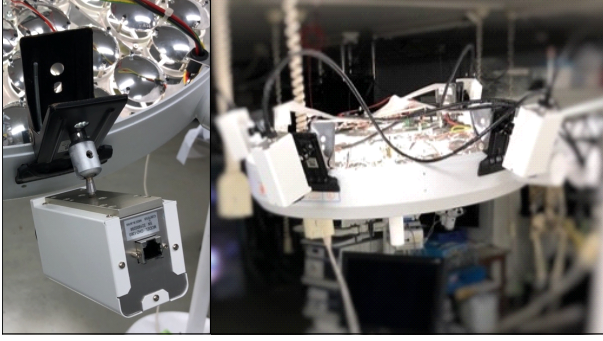


図4 カメラを4台取り付け付けた无影灯
Fig.4 Four cameras are mounted on a shadowless lamp.

以下のように定義する.

$$E(V_{c_1, t_1}, V_{c_2, t_2}) = \begin{cases} 1 & (V_{c_1, t_1} = \text{start}) \\ & \cup((c_1 = c_2) \cap (t_1 + 1 = t_2)) \\ & \cup((c_1 \neq c_2) \cap (t_1 + m = t_2) \\ & \quad \cap (t_2 \leq 2m)) \\ & \cup(V_{c_2, t_2} = \text{end}) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

m はユーザが指定する最小フレームである. 1 のときエッジが結ばれる. またこのとき重み W を以下のように定義する.

$$W(V_{c_1, t_1}, V_{c_2, t_2}) = \begin{cases} s_{t_2}^{c_2} & (V_{c_1, t_1} = \text{start}) \\ & \cup((c_1 = c_2) \\ & \quad \cap (t_1 + 1 = t_2)) \\ \sum_{i=t_1+1}^{t_2} s_i^{c_2} & ((c_1 \neq c_2) \cap (t_1 + m = t_2) \\ & \quad \cap (t_2 \leq 2m)) \\ 0 & V_{c_2, t_2} = \text{end} \end{cases} \quad (4)$$

3.2.2 最適化

節 3.2.1 で生成されるグラフでダイクストラ法を適用し, スタートとエンドノードを除いたノードの情報を取得することで, カメラ番号の配列を取得する. しかし, 最小フレーム先のノードに繋いだエッジが存在するため, その配列は実際の動画のフレーム数より少なくなる. そこで, 配列内のカメラ番号が変わった位置で, そのカメラ番号を飛ばされた分だけ配列に追加する.

4 実験

図4に示すように无影灯に術野カメラを4台設置し, 料理映像での実験を行った. その際, 術野を認識する代わりに, 作業者の手を認識した. 検出器の学習は,

料理映像から無作為に抜き出した約 100 枚の画像を用いて行った. 実験では, 最小フレームを 10 秒間とした. 図5にその結果を示す. グラフ上段からわかるように, カメラの切り替えを制御しなかった場合, カメラの切り替えが頻繁に行われている. しかし, 提案手法を適用した場合, カメラの切り替えが抑えられていることがわかる. また, グラフ中段から見て取れるように, カメラの制限を行った場合でも, そのスコアは限りなく制限を行う前に近いことがわかる. したがって, 手術の観測を維持しながら, QoV を向上させたと言える.

5 むすび

提案手法では, 无影灯に備えられた複数の光源に対応する形で複数台カメラを設置し, カメラの切り替え, 医師の頭部や体を抑えながら手術を記録が可能となった. それによって, 医師は術中にカメラの存在を意識することなく, 手術の記録を行うことができる. また, 献体での実験を通して, その有用性を議論した. 本論文で提案したカメラの切り替え手法は, スコアの算出方法によって, スポーツ等あらゆる方面への応用が期待される. 今後は, カメラの角度や手術工程などを考慮した, 多角的なスコアの算出方法を研究し, より医師の要求にあったカメラ切り替え映像の生成を行いたいと考えている.

謝辞

本研究はAMED 研究費(課題番号 JP18he1902002h 0001), JST CREST(JPMJCR14E1, JPMJCR14E3)の助成を受けたものである.

参考文献

- [1] Kumar A. Sampath, Harinder Pal: Digital video recording of cardiac surgical procedures; The Annals of thoracic surgery, 77(3), 1063-1065 (2004.3)
- [2] Byrd Robert Joe, et al.: Surgical lighting system with integrated digital video camera; U.S. Patent No. 6,633,328. 14 (2003.10)
- [3] Shokei Matumoto, et al.: Digital video recording in trauma surgery using commercially available equipment; Scandinavian journal of trauma, resuscitation and emergency medicine, 21(1), 27 (2013.4).
- [4] John S.K. Murala, et al.: Digital video recording of congenital heart operations with "surgical eye"; The Annals of thoracic surgery, 90(4), 1377-1378 (2010.3)
- [5] Akshay Gopinathan Nair, et al.: Surgeon point-of-view recording: using a high-definition head-mounted video camera in the operating room; Indian journal of ophthalmology, 63(10), 771 (2015.10)
- [6] Jianhui Chen, Peter Carr: Autonomous camera systems: A survey; Workshops at the Twenty-

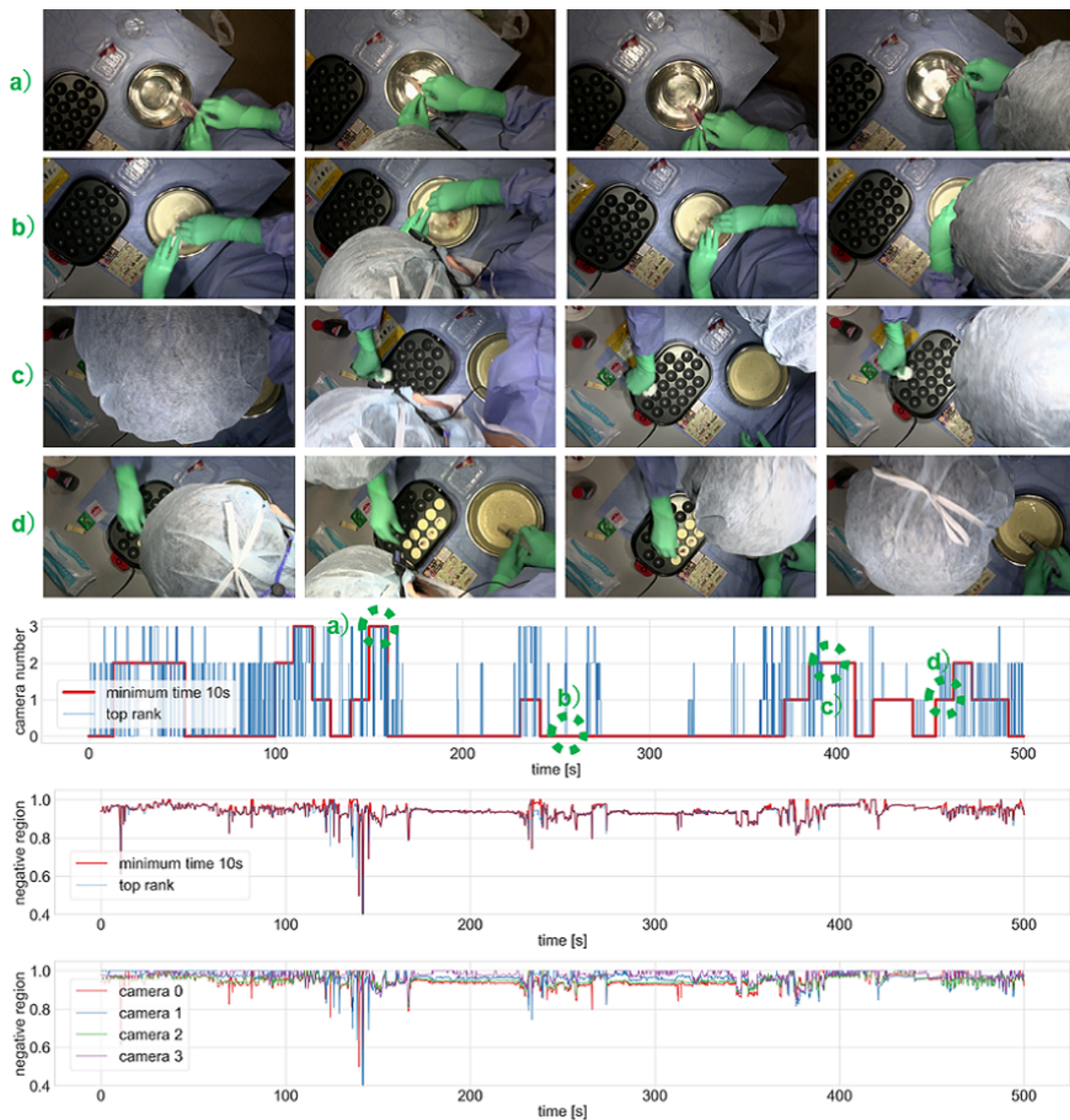


図5 画像例と、提案手法によるカメラスイッチングの結果とスコアのグラフ。; a: 4672 フレームでカメラ 3 が選択されている。b: 7582 フレームでカメラ 0 が選択されている。c: 11704 フレームでカメラ 2 が選択されている。d: 13812 フレームでカメラ 1 が選択されている。; (グラフ上段) カメラ切り替えの結果。(グラフ中段) 最適化後のカメラスコア。(グラフ下段) 各カメラのスコア。

Fig. 5 Example of each camera view, and graphs of camera-switching result and camera scores.; a: Frame 4672 where camera3 is selected. b: Frame 7582 where camera0 is selected. c: Frame 11704 where camera2 is selected. d: Frame 13812 where camera1 is selected.; (Graph upper row) Camera switching result. (Graph middle stage) Score after optimization. (Graph lower row) Score of each camera.

- Eighth AAAI Conference on Artificial Intelligence, 18-22 (2014.7)
- [7] Daniyal Fahad, Andrea Cavallaro: Multi-camera scheduling for video production; 2011 Conference for Visual Media Production(CVMP), 11-20 (2011.11).
 - [8] Daniyal Fahad, Murtaza Taj, Andrea Cavallaro: Content and task-based view selection from multiple video streams; Multimedia tools and applications, 46(2-3), 235-258 (2010.1)
 - [9] Jiang Hao, Sidney Fels, James J. Little: Optimizing multiple object tracking and best view video synthesis; IEEE Transactions on Multimedia, 10(6), 997-1012 (2008.10)
 - [10] Qureshi Faisal Z.,Demetri Terzopoulos: Surveillance in virtual reality: System design and multi-camera control; Computer Vision and Pattern Recognition(CVPR), 1-8 (2007.6)
 - [11] Daniyal Fahad, Murtaza Taj, Andrea Cavallaro: Content-aware ranking of video segments; Distributed Smart Cameras, 2008 Second ACM/IEEE International Conference on Distributed Smart Cameras, 1-9 (2008. 9)
 - [12] Qiong Liu, et al.: Automating camera management for lecture room environments; Proceedings of the SIGCHI conference on Human factors in computing systems (ACM), 442-449 (2001.4)
 - [13] Petr Doubek, et al.: Cinematographic rules applied to a camera network; Omnivis2004: The fifth Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, 17-29 (2004.5)
 - [14] Li Cheng, Kris M. Kitani: Pixel-level hand detection in ego-centric video; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 3570-3577 (2013.12)
 - [15] Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation; International Conference on Medical image computing and computer-assisted intervention, 234-241 (2015.10)
 - [16] Ravindra K. Ahuja, Kurt Mehlhorn, James Orlin, Robert E. Tarjan: Faster algorithms for the shortest path problem; urnal of the ACM (JACM), 37(2), 213-223 (1990.4)