脳神経外科における遠隔手術に適する映像圧縮率とその品質に関する新たな評価法とその結果

New Evaluation Method and Result of Video Compression Ratio and Quality for Telesurgery in Neurosurgery

清水健太郎^{a*}, 村垣善浩^a, 中村亮一^{a,b}, 鈴木孝司^a, 伊関 洋^a

^a 東京女子医科大学大学院 医学研究科 先端生命医科学系専攻

^b 千葉大学大学院 工学研究科 人工システム科学専攻

Kentaro Shimizu^{a*}, Yoshihiro Muragaki^a, Ryoichi Nakamura^{a,b}, Takashi Suzuki^a, Hiroshi Iseki^a

^a Department of Advanced Biomedical Engineering and Science, Graduate School of Medical Science, Tokyo Women's Medical University

^b Division of Artificial Systems Science, Graduate School of Engineering, Chiba University

Abstract

We propose a new evaluation method for video of tele-neurosurgery under microscopic view. We conducted a subjective evaluation of video by 16 evaluators to determine the minimum video bit-rate required to make decisions in tele-neurosurgery under microscopic view. In this evaluation, we developed a new evaluation method that is unaffected by learning from evaluating the same videos more than one. Our results showed that a video bit-rate of at least 384 kbps was necessary in order to make decisions in tele-neurosurgery. Also, contrary to general expectation, the expert subjects scored the low quality bit-rate videos higher than the residents subjects did. Moreover, if the quality of the original picture was poor, the evaluation result was not related to the compression ratio. These finding should be took into account in future evaluation of tele-surgery systems.

Key words

Tele-surgery, Video Bit-rate, H.264 codec.

1. はじめに

医療の発達により、部位や疾患ごとに専門分野が細分化され、より細やかで高度な医療行為が実現した。しかし、たとえば離島などの大都市圏を離れた人口の少ない地域まで、その高度な医療を行う専門医を配置することは、事実上不可能である。これを解決するた

めの一つの方策が遠隔医療であり、宇宙空間や戦場など特殊環境で必然であった遠隔医療のニーズが、医療レベルの向上により一般社会において高まってきているといえる。

近年のめざましいネットワークの発達により、遠隔医療は様々な形で実現されつつある¹⁾⁷⁾。また、遅延などの遠隔医療を実現する際の課題についても検討が進められ、一定の指針が報告されるようになった⁸⁾⁹⁾。しかし、高いネットワーク速度を要求する映像に関しては、大泉らが問題点を指摘しているように²⁾、品質面での課題が残っている。

高い映像品質を確保した通信を実現するためには、

* 東京女子医科大学大学院医学研究科
〒162-8666 東京都新宿区河田町 8-1
kshimizu@abmes.twmu.ac.jp

受付 2008 年 11 月 25 日 ; 採択 2009 年 5 月 11 日

(a)

(b)

Fig. 1 Video Image.

より多くの情報を短時間で伝送できる高速（広帯域）ネットワークを用意することが最も簡単である。たとえば、専用線サービスを用いることで上記の通信環境構築が可能だが、専用線は導入コストが高く、利用できる施設は限られている。したがって、遠隔医療の普及には、低コストで導入可能なネットワーク、つまり、限られた帯域のネットワークであっても、医師の必要とする情報を、常に安定して伝送できることが重要であると言える。

限られた帯域のネットワークで情報を安定して伝送するには、情報を圧縮しなければならない。ただし、情報を圧縮しすぎると品質が劣化し、医療用途には使えなくなってしまう。遠隔読影などのような、静止画を利用するケースについては、JPEG と呼ばれる画像圧縮技術が普及したことで、圧縮率と画像品質の関係に関する多くの研究報告がなされている¹⁰⁾⁻¹²⁾。一方、映像については、たとえば、一般的な分野ではテレビ電話の映像品質を評価する手法¹³⁾や医療分野では救急医療向け映像の品質評価結果¹⁴⁾が報告されている。前者は、基準映像と処理後映像を並べて表示し、両者の違いについて主観により比較評価する手法であり、基準映像に対する評価は可能だが、基準設定できない評価、たとえば、遠隔手術を遂行するに十分な品質かを評価することはできない。後者は、映像信号とノイズ信号の比である PSNR (Peak Signal to Noise Ratio) を用いて処理後映像単独で評価を行ない、さらに裏付けとして専門医による主観評価結果を報告しており、救急医療に十分な品質が得られるかを検討している点で興味深い。しかし、PSNR は、前述の評価方法と同じく、たとえば遠隔手術を遂行するに十分な品質などの基準を満たしているかの評価はできない。また、

裏付け評価を行なう医師が 3 名と少なく、さらに、一人で同じシーンの映像を複数回見るため慣れの影響が無視できず、評価結果には疑問も残る。

そこで本検討では、上述の問題点を回避する方法を取り入れた評価法を用いて映像の品質評価実験を行なったので、その結果を報告する。

2. 対象と方法

本検討では、脳神経外科における遠隔手術で、何らかの判断を下すために最低限必要な圧縮率（以下、ビットレート）を示すことを目標とした。映像圧縮技術の国際標準の一つである H.264¹⁵⁾¹⁶⁾で圧縮した脳神経外科手術の映像を用いて、複数の脳外科医を対象に評価結果に影響する因子を回避できる評価手法を導入し、主観評価実験を行なった。

2.1 実験映像と評価者

実験映像は、脳神経外科の脳動脈瘤手術におけるクリッピング (Fig. 1a) およびその後の手技（たとえば Fig. 1b のような穿刺）のシーンを利用した。クリッピングを対象とした理由は、視野の移動が少ないこと、短時間で重要な判断や手技を行なうポイントが複数あること、である。シーンの選択は、品質評価実験には参加しない熟練医の指導のもと、特に手術中に注意を要する場面、つまり、映像品質が悪いと手術に影響を与える可能性のある場面、を基準として行ない、16 症例を選んだ。これら選択した症例一つにつき 1 本の実験映像を作成することとし、一つの実験映像は、30 ～ 60 秒の長さで、大きさは 720 × 480 ドット、1 秒間に表示するコマ数は、一般のテレビ放送と同じ毎秒約 30 コマとした。また、実験映像中の動脈瘤の大きさは、約 100 × 100 ～ 160 × 150 ドット、クリップの

大きさは、種類や向き、写る角度にもよるが、約 200 × 200 ～ 350 × 350 ドットである。

実験映像の圧縮には、H.264 を用いた。H.264 は、DVD などでも広く普及している MPEG2 に比べ、より高い圧縮率で高品質の映像を生成することが可能とされる映像圧縮技術である。H.264 の圧縮アルゴリズムには、時間的に前後の情報を利用しているものがある。具体的には、映像が時間的に連続しているため隣り合うフレームの相関が高い（似たような図柄である）という特徴を利用し、ある 1 枚のフレームを基準として、それ以降のフレームは基準となるフレームとの差分だけを送るアルゴリズムである。この差分を計算する際に、時間的に未来のフレームの情報を使うことも可能だが、あるフレームを再現するために、さらに先のフレームの情報が届くのを待たなければならないため、遅延を大きくしてしまう原因となる。遅延が大きすぎると、本研究の対象としている遠隔手術には適さなくなるおそれがある。

そこで、実験映像を作成する際の圧縮に用いるフレームは、時間的に前（過去）のものしか使わないこととした。未来のフレームを用いるメリットは、シーン転換（たとえば、TV 放送における本編映像から CM に切り替わる瞬間）があった場合にも高い圧縮率を実現できることだが、シーン転換のほとんどない手術映像では、利用しなくても圧縮効率に大きな影響はない。また、予め圧縮映像を作成する場合、1 度だけ計算する方法（1 パス圧縮）と、2 度計算してより高品質な映像を得る方法（2 パス）があるが、今回はリアルタイム利用であることを考慮し、1 パスにて実験映像を作成した。さらに、ネットワーク伝送が前提であることも考慮し、固定ビットレートによる圧縮とした。

本検討における評価者は、16 名の脳外科医とした。映像品質評価に経験の違いが影響する可能性も考慮に入れ、熟練医 8 名、研修医 8 名の構成とした。

2.2 実験方法

本検討では、映像そのものの品質を検討する目的のため、ネットワーク経由の映像提示は行わず、PC ディスプレイに映像を提示し、それを評価者それぞれの主観で評価した。実験場所は、脳神経外科医局内の打ち合わせスペースとし、光源は一般的な蛍光灯とした。また、PC ディスプレイは、ドットピッチ 0.24mm（24mm が 100 ドットに相当）の 12.1 インチのものをを用い、評価者と PC ディスプレイの距離は、およそ 70cm に設定した。

Table 1 Video group and evaluator group.

	256Kbps	384Kbps	512Kbps	640Kbps
1	A	B	C	D
2	D	A	B	C
3	C	D	A	B
4	B	C	D	A

評価の際、同じ評価者が同じシーンの映像を 2 回以上見ると、最初に見た映像の記憶に 2 回目以降の評価が左右される可能性がある。これを回避するため、静止画の品質評価の実験で提案されている手法 (17) の一部を参考にし、以下のような評価手法を考案した。

まず、16 本の実験映像を 4 本ずつ 4 つのグループ A～D に分割した。さらに、各映像について、H.264 を用いて、256Kbps、384Kbps、512Kbps、640Kbps の 4 種類のビットレートで圧縮した。ビットレートの設定は、実験映像のシーン選択の際と同じ外科医と協議し、決定した。評価者については、4 名ずつ 4 つのグループ 1～4 に分割した。この際、一つのグループは熟練医 2 名と研修医 2 名で構成されるようにした。その上で、Table 1 のような対応をとり、主観評価した。

例えば、グループ 1 の評価者は、256Kbps で圧縮された映像グループ A の 4 本、384Kbps で圧縮された映像グループ B の 4 本、512Kbps で圧縮された映像グループ C の 4 本、640Kbps で圧縮された映像グループ D の 4 本の、合計 16 本を評価する。したがって、例えば 640Kbps で圧縮された映像グループ A を評価することはなく、同じシーンの映像を二度見ることはない。なお、評価者一人につき 16 本の映像を評価するため、全体では 16 人 × 16 本で、256 の評価結果が得られることになる。また、熟練医に対しては、圧縮していない顕微鏡記録映像（以下、オリジナル映像）も評価対象とした。評価にあたっては、評価者にはオリジナル映像であることは知らせず、圧縮映像と同じ条件で評価した。

なお、画面内（1024 × 768 ドット）には、中央に実験映像だけを表示し、それ以外の部分は黒く表示した。

評価方法は、表示された映像で実際に執刀可能か、何らかの判断が可能かを三段階（0：不可能，1：場合によっては可能，2：可能）で行なうこととした。評価結果はビットレートごとに集計し、ビットレート間の評価結果に有意差があるかを統計的に解析した。この実験で得られた評価結果は、正規分布に従うという前提は使えないため、代表的なノンパラメトリック統

	256Kbps	384Kbps	512Kbps	640Kbps
Expert	0.0015	—	—	—
384kbps Resident	0.0377	—	—	—
Total	0.0002	—	—	—
Expert	0.0002	0.498	—	—
512kbps Resident	0.0094	0.488	—	—
Total	0.0001	0.276	—	—
expert	4.96×10^{-4}	0.386	0.792	—
640kbps Resident	5.55×10^{-3}	0.330	0.773	—
Total	1.47×10^{-6}	0.168	0.713	—
Expert	5.74×10^{-4}	0.202	0.485	0.701
Original Resident	—	—	—	—
total	—	—	—	—

Fig. 2 Result of video evaluation (mean value) and P-value.

Table 2 P-value of Expert - Resident.

	Total	256Kbps	384Kbps	512Kbps	640Kbps
P-Value	0.003	0.483	0.063	0.061	0.115

計解析手法であるマン・ホイットニーの U 検定を用いて判定 ($p < 0.05$) した。

3. 結果

映像品質評価の結果、ビットレートが上昇するごとに評価は向上し、また、熟練医の方が、研修医よりも同一品質の映像に対する評価点数が高かった (Fig. 2)。この評価結果を統計解析したところ、256kbps と 384kbps 以上で有意差あり (Fig. 2) となったが、それ以外では有意差はなかった。さらに、各ビットレートとオリジナル映像の評価結果の有意差についても調べ、384kbps 以上でオリジナル映像と有意差なしとなった。熟練医と研修医の評価の違いでは、熟練医の評価点数が研修医よりも高くなった。この差について解析したところ、実験映像全体の評価結果について有

意差ありとなった (Table 2)。

映像ごとの評価では、ビットレートが上がるごとに評価点数が上昇する傾向となったが、オリジナル映像の評価が悪いと、一番高いビットレートの 640kbps であっても意志決定不可能と評価されるものもあるなど、圧縮した映像の評価結果は前述の傾向とは一致しなかった (Table 3)。

4. 考察

ビットレートと映像品質の関係について、384kbps 以上で評価結果に有意差なしとなったことから、H.264 で圧縮した映像では、手術中の意志決定には最低限 384kbps のビットレートが必要であると言える。Kim らが救急医療向けの映像品質について検討結果を提示している¹⁴⁾が、一人の評価者が同じシーンの映像を複数回見るため、評価結果には複数回見ることによる記憶のバイアスがかかってしまうことが否めない。本検討では、一人の外科医が同じシーンの映像を複数回見ることはない新しい評価手法を構築したことで、こ

Table 3 Result of evaluation (Each Video).

	256Kbps	384Kbps	512Kbps	640Kbps	Original
1	0.75	1.5	1.75	2	2
2	0.75	1.5	1	1.5	2
3	1	1.5	1	2	2
4	0.75	2	1.75	2	2
5	0.25	1.25	1	1	2
6	0.5	1.75	2	1.75	2
7	0.75	1.5	1	1.75	2
8	2	1.25	2	1.5	2
9	0.5	0.25	1	0.5	2
10	1	1	1.5	1	1.5
11	0	0.5	1	1.25	1.5
12	0.25	1.25	1.25	1.75	1.5
13	0	0.5	0	0.5	1.5
14	0	1	1.25	0.75	1
15	0	1	1.75	1.5	1
16	0.25	0.25	1.5	0.75	0

の問題点を解決しているため、記憶によるバイアスを回避した評価結果が得られたと言える。

熟練医と研修医の評価結果の違いについては、実験映像全体の評価結果で有意差が認められた。実験前は、熟練医の方が高品質な映像を要求すると考え、厳しい評価を下すと予想していたが、ここで示された結果は予想とは逆となった。実験映像全体で評価結果に有意差が生じた理由としては、熟練医は、多くの手術経験から解剖学的な知識も豊富であり、多少の情報の欠落については、その知識を用いて補完できるためと考えられる。研修医については、経験や知識が少なく、映像で提示された情報がすべてとなるため、微少な映像の乱れであっても対応できず、評価が低くなっていると考えられる。ビットレートごとでは有意差はないものの、256kbpsと384kbpsではP値が大きく異なり、熟練医と研修医の差が広がっていると言える。実験中、外科医は、クリップをかける場所を中心に、目的の位置にクリップがかかっているか、周辺にある別の血管まで挟み込んでいないか、に注目していた。256kbpsでは、これら注目点が熟練医の経験や知識でも補えないほどに失われてしまい、研修医との差がなくなったが、384kbpsでは熟練医であれば補完が可能なレベルで情報が残ったため、差が開いたと考えられる。ただし、実験映像に対する評価は、評価者個人の評価傾向の違いも影響するため、厳密には熟練度の違いによるものか評価傾向の違いによるものかの区別はできない。より詳細な検討をするためには、評価者の人数をさらに増やし、評価者個人の評価傾向の影響を受けに

Fig.3 Sample of indistinct video image.

くくすることが重要である。

映像ごとの評価結果では、一部の映像ではビットレートが上昇しても評価が低い（1点以下）ままで、中にはビットレートの上昇とは逆に評価が下がっているものがある（たとえばTable 3における映像16）。この現象は、オリジナル映像の評価点数が低い場合に多く発生していた。評価が低いオリジナル映像を調査したところ、映像のフォーカスが他の映像とくらべて甘く、全体的にややぼやけていた（たとえばFig. 3）。フォーカスの甘さから、意志決定に必要な情報を得ることが困難なために全体的な評価が落ちるとともに、判断基準が曖昧となってビットレートの上昇と関係ない評価点数の動きになったと考えられる。また、ビットレートの上昇とともに評価点数が上昇しない事例はオリジナル映像の評価が良い映像でも存在した（たとえば映像6）。これは、映像1本あたりの評価人数が4人と少ないため、評価のバラツキが現われたものと考えられる。ただ、16本の映像全体では評価のバラツキはないため、今回の検討目的には影響はないと考えられる。映像ごとの評価を詳細に検討するためには、映像1本あたりの評価人数を増やす必要がある。方法としては、評価者の人数を32名にすること、あるいは、評価するビットレートの種類を2種類にした上で、評価者と映像のグループを2グループにすることで、映像1本あたりの評価人数を倍の8人にできる。これらの点を考慮した実験系を構築することで、われわれの提案した品質評価手法は、有効な評価結果を得ることができると考えられる。

5. 結論

本検討では、遠隔手術に適した映像品質について、

一定の指針を示すことを目標とし、遠隔手術用映像の品質を評価するための新しい手法を提案し、実験を行った。その結果、遠隔手術において何らかの意志決定をするためには、映像のビットレートは最低限 384kbps 必要であることを示した。また、熟練医と研修医の評価差は、経験による解剖学的な知識の差から、当初の予想とは逆に、熟練医の方が映像のビットレートが低くても評価が良いことがわかった。さらに、オリジナル映像の品質が悪い場合、評価結果はビットレートの上昇と関連しなくなることも確認された。これらの結果は、遠隔手術システムを構築する上で重要な示唆であり、われわれの提案手法の有効性が示された。

謝辞：実験データの提供、作成、および評価実験にご協力くださった、東京女子医科大学病院脳神経外科の堀智勝教授、岡田芳和先生、丸山隆志先生、佐々木寿之先生、川島明次先生、河本竹正先生、米山琢先生、山口浩司先生、井上龍也先生、中嶋剛先生に、心より感謝申し上げます。また、評価実験にご協力くださった、中村記念病院の及川光照先生に、心より感謝申し上げます。

本研究は、NEDO 研究開発事業「基礎研究から臨床研究への橋渡し促進技術開発／橋渡し促進技術開発疾患動物を用いた新規治療機器の安全性・有効性評価手法の開発」、および、柏森情報科学振興財団研究助成「術中情報可視化・統合・解析支援システム「手術戦略デスク」の開発」(K17 研 X 第 233 号) の助成を受けている。

文献

- 1) 日本顎関節学会. 顎関節疾患および顎関節症の分類. 日本顎関節学会誌 1996 ; 8(2): 112-117.
- 2) 飯塚忠彦, 大西正俊, 井上宏, 覚道健治, 村上健一郎, 大月佳代子. 顎関節症診断治療マニュアル. 永永書店, 2004 : 2-7.
- 3) NIH, U.S. Department of Health and Human Services. TMJ Disorders. NIH Publication No. 06-3487, 2006.
- 4) 羽毛田匡. 顎関節症と開口障害 顎関節症のスプリント療法. ENTONI 2005 ; 51 : 44-51.
- 5) 五十嵐千浪, 小林馨. 顎関節症 筋肉に痛みがある場合にはどのようにすりを使いますか?. 歯界展望別冊 2007 : 147-148.
- 6) DeBar LL, Vuckovic N, Schneider J, Ritenbaugh C. Use of complementary and alternative medicine for temporomandibular disorders. Journal of Orofacial Pain 2003;17(3):224-236.
- 7) 橋本修一. 顎関節症への理学療法アプローチ. 理学療法ジャーナル 2000 ; 34(4) : 270-271.
- 8) 齊藤一郎. ドライマウスの臨床. 医歯薬出版, 2007.
- 9) NIH, U.S. Department of Health and Human Services. Dry mouth. NIH Publication No. 06-3174. 2006.
- 10) Saunders JR, Hirata RM, Jaques DA. Salivary glands. Surgical Clinics of North America 1986; 66(1):59-81.
- 11) 阪井丘芳. 唾液腺・口腔粘膜マッサージ～唾液の機能と口腔ケア. 日本歯科医師会雑誌 2008 ; 60(12) : 1289-1292.
- 12) Koga H, Usuda Y, Matsuno M, Ogura Y, Ishii H, Solis J, Takanishi A, Katsumata A. Development of Oral Rehabilitation Robot for Massage Therapy. Proceedings of the International Special Topic Conference on Information Technology in Biomedicine 2007, 2007:111-114.
- 13) Koga H, Usuda Y, Matsuno M, Ogura Y, Ishii H, Solis J, Takanishi A, Katsumata A. Development of the Oral Rehabilitation Robot WAO-1. Proceedings of the 2nd Biennial IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechanics, 2008.
- 14) Salisbury JK. Active Stiffness Control of a Manipulator in Cartesian Coordinates. Proceedings of the 19th IEEE Conference on Decision and Control, 1980: 95-100.
- 15) 杉本浩一. ロボットアームのカフィードバック制御. 計測と制御 1986 ; 25(1) : 45-50.
- 16) Sugahara Y, Hosobata T, Mikuriya Y, Lim H, Takanishi A. Realization of Stable Dynamic Walking by a Parallel Bipedal Locomotor on Uneven Terrain Using a Virtual Compliance Control. Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003: 595-600.
- 17) 平林久明, 杉本浩一, 荒井信一, 坂上志之. 多自由度ロボットの仮想コンプライアンス制御. 計測自動制御学会論文集 1986 ; 22(3) : 343-350.
- 18) 小林馨. 顎関節症のみかた. 永末書店, 2004.
- 19) 勝又明敏, 飯田幸弘, 藤下正巳, 泉雅浩, 有地淑子, 有地栄一郎, 小木信美, 五十嵐千浪, 小林馨, 古賀裕樹, 石井裕之, 高西淳夫. 咀嚼筋の徒手のマッサージによる開口度改善. 日本顎関節学会誌 2008 ; 20 : 91.
- 20) 齊藤一郎. ドライマウス. 日本評論社, 2003 ; 160.
- 21) 小島麻衣子, 星野里衣子, 高井尚子, 加藤節子. 口腔内自浄作用を高めるために - 唾液腺マッサージによる口腔内環境の改善を試みる. 日本看護学会論文集, 老年看護 2003 ; 34 : 159-161.