

動画画像スケーラブル変換制御におけるクロップ領域条件に関する研究

藤村 裕代[†] 花村 剛[†] 市川 正浩[†] 伊藤 朝香[†]

本研究は、受信端末の提示能力や動画画像コンテンツの内容等に応じて映像をスケーラブルに変換し、視聴者に配信するためのアルゴリズムの確立を目的とする。本実験では特に、HD サイズの動画画像を QVGA 及び QQVGA サイズに変換する際の最適なクロップ(切り取り)領域を明確にすることを目的とし、人間工学に基づいた評価実験を行った。その結果、人が動画画像から受ける印象や評価はクロップ領域の条件によって大きく異なること、動画画像の属するカテゴリによって適切なクロップ領域は異なることが分かり、その最適条件を絞り込んだ。また視聴時の眼球運動から、特に QQVGA サイズの動画画像に関して、注視点や瞬目、瞳孔面積の特性や主観評価との関連性について、いくつかの傾向が示唆された。

A Study on Video Scalable Conversion Control System according to the Condition of Cropped Range

Hiroyo FUJIMURA[†] Tsuyoshi HANAMURA[†] Masahiro ICHIKAWA[†] and Asaka ITO[†]

This research establishes an algorithm to create scalable video contents according to the video contents or the display capacity of the receiving terminal, and deliver it to the user. The purpose of this experiment was to find the appropriate crop (cut) area when converting HD videos to QVGA and QQVGA size. We observed that evaluation or impact received by the user varied greatly depending on the crop area. Also, the size of the appropriate crop area varied according to the content and category of the video. So the crop area was narrowed down to optimum conditions. Especially when converting to QQVGA size, we found that blinking, fixation point and pupil size might be connected with subjective evaluation.

1. まえがき

放送番組コンテンツの提示に対して視聴者が要望する項目として、内容の充実とともに高精細な表示があげられている。しかし、現状のように大画面から携帯端末のような小さな画面に至るまで、多様な表示形態がある場合には、各々の受信端末に合わせて個別に番組制作を行うのは効率的でなく、それらの提示能力や放送番組コンテンツの内容に応じて、送出側または受信側で自動的に動画画像をスケーラブルに変換する技術が求められる。

そこで本研究では、HD サイズ(1920x1080 pixel)の動画画像を今後受信端末としての利用が想定される携帯端末に適したサイズに処理する際の最適なスケーラブル変換制御アルゴリズムについて、人間工学に基づき評価し、確立する

ことを目的としている。今回はクロップ領域の違いが人に与える影響について、評価検討を行った。以下にそれを報告する。

2. スケーラブル変換制御のアルゴリズム

HD サイズの映像を QVGA(320x240pixel)や QQVGA(160x120pixel)サイズに変換する場合、その処理方法によっては、被写体の識別が困難となる、あるいは変換前後で視聴者に与える印象が大幅に変化する、といった問題が発生する。そこで、最適なスケーラブル変換の方式を確立するためには、これらの問題の原因と考えられる項目について検討し、変換フローの構成要素として適切な条件を設定することが必要だと考えられた。以下に構成要素とした事柄を列挙する。

[†] 株式会社 メディアグルー 技術統括部
R&D Department, Media Glue Corporation

- (a) 抽出するオブジェクトまたはポイント
(主役となる物体：Leading Part。以下、LP)
- (b) LP 照準点
- (c) クロップ領域
- (d) ダウンスケール方法
- (e) オブジェクトの特殊加工

ここで、(a)のLPとは、一連の変換処理の基準となる映像内のオブジェクトもしくは点とし、映像の中央、動きベクトルの大きいオブジェクト、といった条件の設定である。(b) LP 照準点とは、変換後の画面におけるLPの位置である。設定条件として、画面の中央、垂直・水平方向を各々3分割した交点、変換前後のX,Y座標と同比率の位置とするなどが挙げられる。(c)は動画像上の変換処理対象領域の指定、(d)は変換時の縮小方法であり、全変換対象領域を均一に縮小、あるいは周囲部の大幅に縮小し中央部は縮小度合いを抑えるといった設定条件が挙げられる。(e)は色や輝度の変更、コンテンツ再生時間の短縮等の設定である。

以上の構成要素のうち、(c)クロップ領域の指定は、図1に示すように変換後の動画像に特に大きな変化を及ぼすため、動画像のサイズを大幅に小さくする変換の際には、とりわけ重要な要素となる。

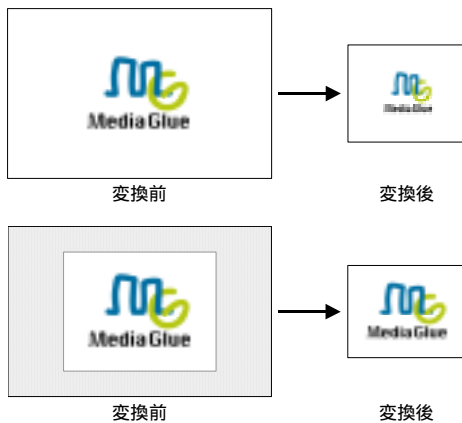


図1 クロップ領域の違いと変換後の映像
(上:全画面を縮小(スクイーズ))
(下:点範囲以外をクロップ領域として縮小)

3. 評価実験

3.1 目的

本実験では、変換フローの構成要素のうち、特に(c)クロップ領域に関して、HDサイズの映像をQVGA及びQQVGAサイズに変換する際の最適条件を明確化することを目的として、人間工

学に基づく主観評価により、設定条件の絞り込みを行った(3.2参照)。更に、動画像視聴時の眼球運動を測定し、クロップ領域の違いが人に与える影響について、客観的側面からも検討を行った。(3.3参照)

3.2 評価実験

3.2.1 実験条件

クロップ領域は1動画像に対し、変換後のサイズがQVGAとなる場合は計8通り、QQVGAとなる場合はQVGAの条件に加え480x320pixelサイズを含む計9通りとした。クロップ領域の水平・垂直方向の比率は、変換後の画面比率と等倍とし、画面アスペクト比4対3とした。クロップ領域以外の構成要素の条件は、(a)動画像の中央をLPに指定、(b)LP照準点を画面中央、(d)全クロップ領域を均一に縮小、(e)未加工、で固定とした。

表1 評価を行ったクロップ領域

クロップ領域 (pixel)	480	640	800	960	1120	1280
クロップ領域:HD	3:12	4:12	5:12	6:12	7:12	8:12

pixel値(上列)、HDとのサイズ比(下列)共に水平方向クロップ領域の480は、QQVGAサイズのみの変換条件

また、HDサイズからSDサイズに変換する際の既存方式である、レターボックス、エッジクロック、スクイーズの3手法も検討条件に加え評価を行った。

実験には計18種類、各20秒間の動画像を用い、上記の条件にて変換処理したものを呈示刺激とした。なお、動画像はBSデジタル放送推進協会のカテゴリに基づき、「紀行・情報・ドキュメンタリ」、「ドラマ」、「映画」、「音楽・舞台」、「スポーツ」、「クイズ・バラエティ」の6つのカテゴリに3種ずつ分類された(表2参照)。

表2 評価動画像の内容とカテゴリ分類

カテゴリ	コンテンツ内容
紀行・情報・ドキュメンタリ	料理番組、リゾート地紹介、ニュース
ドラマ	ドラマ紹介、時代劇、サスペンスドラマ
映画	洋画1、洋画2、洋画紹介
音楽・舞台	オーケストラ演奏、ミュージック番組、宝塚
スポーツ	アイスホッケー、スピードスケート、相撲
クイズ・バラエティ	クイズ、トレンド情報、幼児向け番組

また、上記 18 動画像のうち、料理番組、ニュース、クイズ、トレンド情報及びカテゴリ「映画」の 3 動画像は字幕付きである。

3.2.2 評価方法

20 歳代の男女 8 名を被験者とし、呈示刺激である全 126 種の QVGA サイズの動画像及び全 162 種の QQVGA サイズの動画像について評価を求めた。動画像はランダムに呈示し、テロップなどの文字情報についても評価基準に含め、非常に好ましいを 5 点、やや好ましいを 4 点、どちらともいえないを 3 点、やや好ましくないを 2 点、非常に好ましくないを 1 点とする 5 件法での回答を求めた。

3.2.3 結果

(1) QVGA サイズ

クロップ領域を要因とする分散分析では 1% 水準で有意差が認められ ($F=96.972, P<0.0001$)、レターボックス、エッジクロック、スクイーズは総じて評価が高く、なかでもレターボックスの評定点が最も高くなった。クロップ領域条件を 1280x960pixel とした動画像の評定点も高くなり、領域の大きさと評定点は比例関係に近い傾向があった。

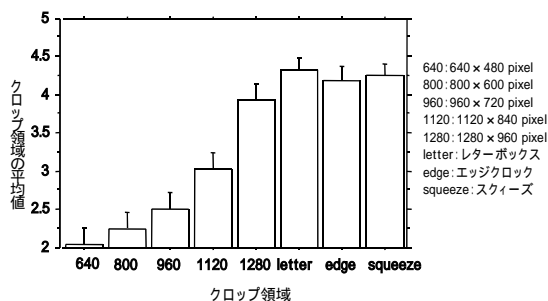


図 2 クロップ領域別評価点(QVGA サイズ)

カテゴリを要因とした 1 要因の分散分析を行ったところ、1% 水準で有意差が認められた ($F=9.070, P<0.0001$)。「ドラマ」及び「スポーツ」は他のカテゴリと比較し評価が高く、反対に「紀行・情報・ドキュメンタリ」、「クイズ・バラエティ」は有意に低い結果が得られた。なお、同カテゴリに属する動画像間の差は、「音楽・舞台」が最も大きく、また高い評価を得たカテゴリである「スポーツ」の 3 動画像はいずれも高い評定点となった。

カテゴリとクロップ領域を要因とした、6×8 の 2 要因の分散分析では、1% 水準で有意差が認

められた ($F=4.525, P<0.0001$)。レターボックス、エッジクロック、スクイーズの評定点はいずれのカテゴリにおいても評価が高く、また、水平・垂直方向のクロップ領域ともに元映像を下回るサイズとした場合、その領域が大きいほど、評定点が高い傾向が見られた。カテゴリ「ドラマ」及び「クイズ・バラエティ」はクロップ領域の違いによる差が顕著に見られ、反対に「スポーツ」では、その差が比較的小さくなった。また、「紀行・情報・ドキュメンタリ」のエッジクロックの評定点は他カテゴリに比べ低く、「映画」ではクロップ領域 1280x960pixel の評定点が他カテゴリの同クロップ領域における結果に比べ、評価が高くなった。

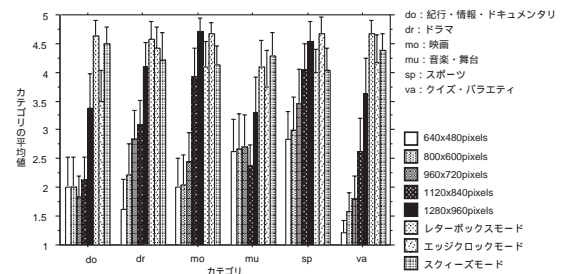


図 3 カテゴリ及びクロップ領域別評価(QVGA サイズ)

なお、字幕の有無を要因とした 1 要因の分散分析では、5% 水準で有意な傾向が認められた ($F=8.575, P=0.0035$)。特に、クロップ領域が 800x600pixel 及び 960x720pixel の評定点に大きな差が見られた。

(2) QQVGA サイズ

クロップ領域を要因とした分散分析では 1% 水準で有意差が認められた ($F=9.070, P<0.0001$)。QVGA サイズの結果と同様に、水平・垂直方向のクロップ領域ともに元映像を下回るサイズとした場合、その領域が大きいければ大きいほど、評定点が高くなった。

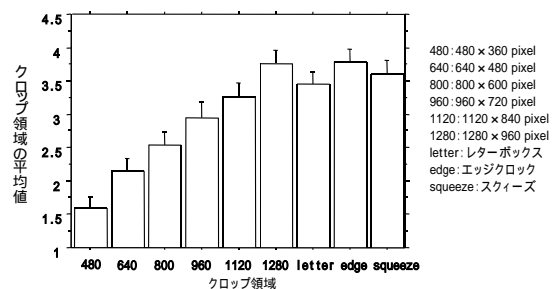


図 4 クロップ領域別評価点(QQVGA サイズ)

カテゴリを要因とした1要因の分散分析を行った結果、1%水準で有意差が認められた ($F=5.821, P<0.0001$)。「ドラマ」及び「スポーツ」は他カテゴリと比較して評価が高く、反対に「紀行・情報・ドキュメンタリ」、「クイズ・バラエティ」はQVGAサイズ同様、評価が低くなった。また、カテゴリ「クイズ・バラエティ」に属する動画間の評価には大きな差が見られた。

カテゴリとクロップ領域を要因とした、 6×8 の2要因の分散分析の結果、1%水準で有意差が認められ ($F=2.623, P<0.0001$)、いずれのカテゴリにおいても、クロップ領域が大きいほど評定点が上昇する傾向が見られた。「映画」は他カテゴリと比較し、クロップ領域を1280x960pixelとした評価がとりわけ高く、レターボックス、エッジロック、スクィーズは低い評定点であった。また、いずれのカテゴリにおいても、クロップ領域の違いにより生じる評定点の高低差はQVGAサイズより小さい傾向が見られた。

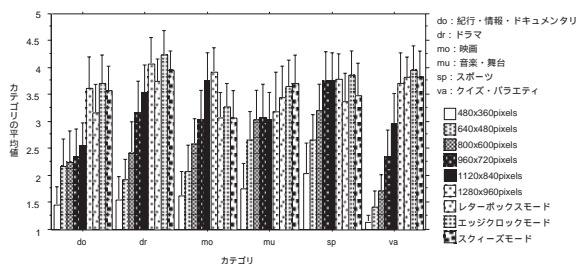


図5 カテゴリ及びクロップ領域別評価(QQVGAサイズ)

なお、字幕の有無を要因とした1要因の分散分析では、1%水準で有意差が認められた ($F=27.304, P<0.0001$)。QVGAサイズと同様、クロップ領域が800x600pixel及び960x720pixelの評定点に大きな差が見られたことに加え、レターボックス、エッジロックに関しても、有意な差が見られた。

3.2.4 考察

評価実験の結果から以下の4点が考察としてあげられた。

- (a) クロップ領域の大きさは、人の主観的な評価に影響を与え、動画コンテンツの一部分をクロップ領域とする場合、その領域の大きさと視聴者の評価は比例関係に近似していた。
- (b) レターボックス、エッジロック、スクィーズの評価は総じて高く、この傾向は変換

後のサイズがより大きい場合に顕著に現れた。

- (c) 変換後のサイズ、及び動画の内容や属するカテゴリによって、適切なクロップ領域が異なっていた。
- (d) 字幕付きの動画は、字幕無しよりも評価が低く、この傾向はクロップ領域の指定等で字幕が途切れてしまう場合、顕著に現れた。また変換後のサイズが小さいほど、字幕の有無によって評価に違いが現れた。

なお、本実験結果から変換フロー構成要素の一つであるクロップ領域の最適条件はカテゴリ毎に表3のように示唆された。

表3 クロップ領域最適条件

カテゴリ	QVGA サイズ	QQVGA サイズ
紀行 情報 ドキュメンタリ	レターボックスモード スクィーズモード	1280x960pixels エッジロックモード
ドラマ	レターボックスモード エッジロックモード	1280x960pixels エッジロックモード
映画	1280x960pixels エッジロックモード	1280x960pixels 1120x840pixels
音楽 舞台	レターボックスモード スクィーズモード	エッジロックモード スクィーズモード
スポーツ	1280x960pixels エッジロックモード	1280x960pixels エッジロックモード
クイズ バラエティ	レターボックスモード スクィーズモード	エッジロックモード スクィーズモード

3.3 評価実験

3.3.1 実験条件

3.2 評価実験 と同様の6カテゴリ18種類の動画をを用い、QVGA、及びQQVGAサイズに変換した動画視聴時の眼球運動を測定した。変換時のクロップ領域は3.2 評価実験 で絞り込んだ表3に示す条件で設定し、変換フローの他構成要素は3.2 評価実験 と同一条件とした。

3.3.2 評価方法

PCモニター中央にQVGA、及びQQVGAサイズに変換した72の動画を呈示刺激として10秒間ずつランダムに表示し、その間、被験者の眼球運動を測定した。なお、被験者は20歳代の男女2名とし、各動画視聴後に、その画質について5件法で回答を求めた。

3.3.3 結果

(1) QQVGA サイズ

5件法での回答の結果、高い評定点を得た動画と低い評定点であった動画視聴時の眼球

運動測定データを比較したところ、以下の結果が得られた。

(1-1) 注視点について

呈示画面の水平方向をX軸、垂直方向をY軸、回数をZ軸とした3次元グラフにおいて、評定点の高い動画視聴時の測定データは、比較的底面が小さく、頂点が底面の中部にある錐体に近い形となる傾向が見られた。反対に評定点の低い動画を視聴時の場合、頂点が底面範囲の端、あるいは範囲外に存在するなどといった、歪みのある錐体となる、あるいは広範囲に注視点がばらついている傾向が見られた。

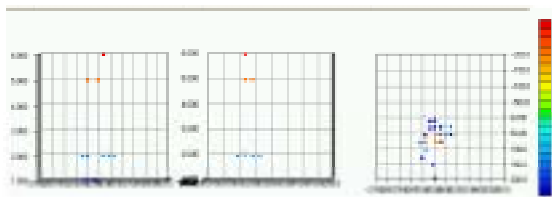


図6 評価の高い映像呈示時の注視点分布図
(右:x画面水平方向 y 頻度 中央:x画面垂直方向 y 頻度 左:x画面水平方向 y 画面垂直方向)

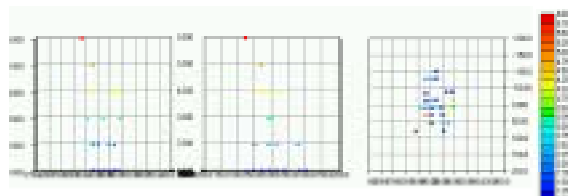


図7 評価の低い映像呈示時の注視点分布図
(右:x画面水平方向 y 頻度 中央:x画面垂直方向 y 頻度 左:x画面水平方向 y 画面垂直方向)

(1-2) 瞬目について

評定点の高い動画視聴時は、評定点の低い動画に比べ、瞬目回数が少ない傾向が見られた。また、評定点が高い場合、視聴開始直後、終了間際に瞬目しているケースが多く、視聴中間である、視聴開始よりおよそ3~7秒間での瞬目は比較的少ない傾向が見られた。(図8参照)

評定点の低い動画視聴時は、再生開始からの時間に関わらず瞬目が比較的多く発生する傾向が見られた。(図9参照)

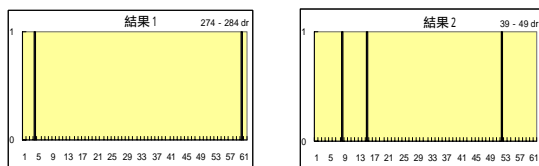


図8 評価の高い映像視聴時の瞬目頻度

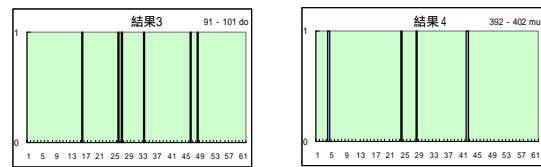


図9 評価の低い映像視聴時の瞬目頻度

(1-3) 瞳孔面積について

瞬目を除き、評定点の高い動画視聴時は、評定点の低い動画に比べ、瞳孔面積の変化が比較的大きく、広範囲で推移している傾向が見られた。(図10,11参照)

また、評定点の高い動画視聴時は、図10左に示すように、その瞳孔面積の最大値は大きな値となる傾向が多く見られた。反対に、評定点が低い場合、その値は小さくなる傾向が多く見られた。

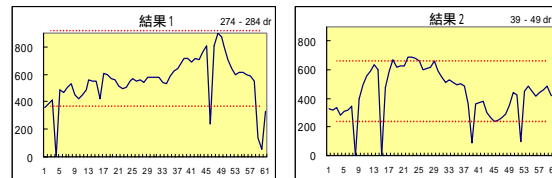


図10 評価の高い映像視聴時の瞳孔面積推移

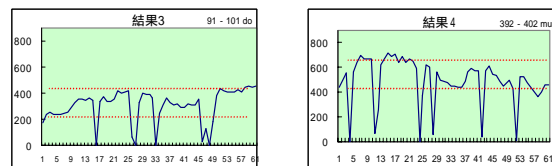


図11 評価の低い映像視聴時の瞳孔面積推移
図10,11の赤点線は瞳孔面積の最大・最小値

(2) QVGA サイズ

QQVGA サイズに比べ、高い評定点を得た動画と低い評定点の動画視聴時の眼球運動測定データには大きな差が見られなかった。

3.3.4 考察

本評価実験より、眼球運動を指標としたQQVGAサイズの評価結果として、主観評価の高い動画視聴時には、以下の傾向が見られた。

- (a) 視聴者は画面内のある一定範囲に集中して注目する傾向が見られた。また、最も頻繁に注視する一点を中心に、頻繁に注目する部分が放射線状に近いかたちで広がっている傾向が見られた。
- (b) 瞬目回数は評価の低い動画視聴時と比較して少ない傾向が見られ、瞬目が見られる

のは比較的動画像を再生した直後に多く、中盤には瞬目回数が減る傾向が見られた。この現象は視聴者が映像に集中するために見られると考えられた。

- (c)瞳孔面積は、その最大と最小値の差が大きく、推移変化量が大きい傾向が見られ、また最大値に関しては、主観評価の低い動画像視聴時の値と比べ、大きい値を示す傾向が見られた。

主観評価の低い動画像視聴時には、以下の傾向が見られた。

- (d)画面上の広範囲に注視点がばらつく傾向が見られ、視聴者の注目する部分が定まっていなかったと考えられた。
- (e)映像再生時間に関わらず、主観評価の高い映像視聴時に比べて瞬目頻度が比較的多い傾向が見られた。
- (f)瞳孔面積の最大・最小値の差は主観評価の高い映像視聴時に比べて比較的少なく、推移変化量は小さい傾向が見られた。また、瞳孔面積の最大値も小さい値を示す傾向が見られた。

4. むすび

本研究では、スケーラブル変換した動画像が人に与える影響についての評価を行った。結果として、クロップ領域の条件によって人が動画像から受ける印象や評価は大きく異なり、元映像の一部分をクロップ領域として指定し変換する場合、その領域が大きいほど評価が高いこと、また、レターボックス、エッジクロック、スクイーズは総じて評価が高いことが明らかになった。更に、動画像の内容やカテゴリによって最適なクロップ領域は異なることが判明し、カテゴリ毎にクロップ領域の最適条件を示唆できた(表3参照)。また、動画像の評価には差が見られたことから、携帯端末への配信に適した動画像やカテゴリが存在すると考えられた。

また、スケーラブル変換を行った動画像を視聴した際の眼球運動について、その特性や主観評価との関連性を検討した結果、特に QQVGA サイズに変換した動画像において、主観評価の高い動画像は、評価の低い動画像視聴時と比較して、注視点は一点に定まっている傾向が見られた。また、瞬目は比較的少なく、瞳孔面積は変化に富み、最大値も大きい傾向も見られた。

今回は、クロップ領域のみに限定した最適条件の絞り込みを行った。今後は、本評価実験の結果をもとに、ユーザの属性や嗜好といった点も考慮しながら、他要素の最適条件についても検討を行う予定である。また、今回は PC モニタを使用した評価実験であったが、今後は携帯電話や PDA 等を用いて、実際の視聴環境に近い状態での評価を行い、最適なスケーラブル変換のアルゴリズムを検証する。なお、人間工学に基づいた主観的評価と客観的評価との関連やその特性についての検討も継続して行い、両側面からスケーラブル変換方式や動画像コンテンツの評価方式を確立する予定である。

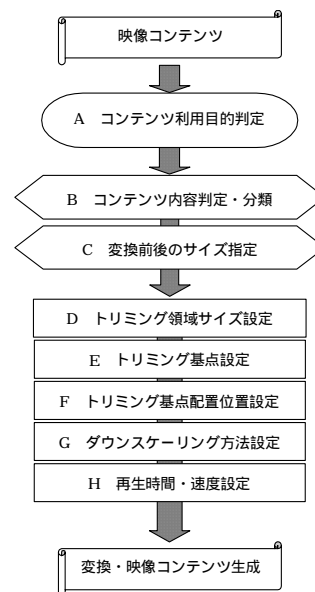


図 12 スケーラブル変換フロー

本研究は通信・放送機構委託研究「次世代放送方式技術の研究開発」の一貫として行っているものである。

参考文献

- [1] 藤原、磯部: "多様な視聴者にやさしい高度放送受信ナビゲーションシステムの検討課題", 映像学冬季大会, 3-1, p66(2001)
- [2] 花村、大和久、市川: "放送動画コンテンツのスケーラブル変換制御に関する検討", 信学全大, D-11-89(2003)
- [3] 大和久、花村、市川: "受信端末に応じた動画像スケーラブル変換制御方式の検討", 映像学技報, Vol.27, 24, pp9-12(2003)
- [4] 泉 武博(監修)NHK放送技術研究所(編): 3次元映像の基礎, オーム社
- [5] 村井、中山、清水: "テレビ番組視聴時の瞳孔面積変化と主観評価の創刊について", 信学技報, ET-98-32, pp7-14(1998)