

多様な視聴スタイルを可能にするコンテンツ変換サービス

宇和田 弘美[†] 灘本 明代[‡] 熊本 忠彦[‡] 濱辺 徹[†] 横澤 誠^{†,††} 田中 克己^{‡,††}

[†] 野村総合研究所 〒240-0005 横浜市保土ヶ谷区神戸町 134

[‡] 情報通信研究機構けいはんな情報通信融合研究センター 〒619-0289 京都府相楽郡精華町光台 3-5

^{††} 京都大学大学院情報学研究所 〒606-8501 京都市左京区吉田本町

E-mail: [†] {h-uwada, t-hamabe, m-yokozawa}@nri.co.jp, [‡] {nadamoto, kuma}@nict.go.jp

^{††} ktanaka@i.kyoto-u.ac.jp

あらまし ネットワーク技術の発展に伴って、コンテンツの流通チャネルが放送と通信の垣根を乗り越えはじめた。コンテンツを配信する経路が増加する一方で、ユビキタス端末の普及により視聴スタイルも多様化する。特定の視聴スタイルを対象としてコンテンツを制作するだけに留まらず、様々な視聴スタイルに応じたコンテンツ制作が求められる。しかしながら、コンテンツの供給者があらゆる視聴スタイルを織り込むことは難しい。このため、供給者と利用者（ユーザ）とを繋ぐコンテンツ変換サービスが求められる。我々は、視聴スタイルと端末との関係を整理し、端末ユーザにとって望ましいコンテンツ変換サービスを提案する。

キーワード 視聴スタイル, コンテンツ変換, ユビキタス

Content Transformation Service for Diverse Audience Styles

Hiromi UWADA[†] Akiyo NADAMOTO[‡] Tadahiko KUMAMOTO[‡] Toru HAMABE[†]

Makoto YOKOZAWA^{†,††} and Katsumi TANAKA^{‡,††}

[†] Nomura Research Institute, Ltd. 134 Godo-cho, Hodogaya-ku, Yokohama, 240-0005 Japan

[‡] National Institute of Information and Communications Technology (NICT) 3-5 Hikari-dai Seika-cho Souraku-gun, Kyoto, 619-0289 Japan

^{††} Graduate School of Informatics, Kyoto University

E-mail: [†] {h-uwada, t-hamabe, m-yokozawa}@nri.co.jp, [‡] {nadamoto, kuma}@nict.go.jp

^{††} ktanaka@i.kyoto-u.ac.jp

Abstract According to the development of network technology, the distribution channel of content crosses over the boundary between broadcast and communication. One side, the delivery channel increases. The other side, the audience style will diverse because of the spread of ubiquitous terminals. An information provider has been designed content targeting for specific style. In the future, the information provider will be required to develop the production methods for satisfying various audience styles. However, it is difficult for an information provider to design content suiting for audience styles completely. Therefore, content transformation service that ties information providers and users, is needed. We put the relationship between audience styles and terminal types into order, and introduce a desirable transformation method for terminal users.

Keyword audience style, content transformation, ubiquitous

1. はじめに

モバイル端末の普及により、大小様々な機器を通してコンテンツが提供されるようになった。インターネットのブロードバンド化とともに、ユビキタス化が進んでいるが、更に我が国では放送のデジタル化も進んでいる。この延長線上に、通信と放送のコンテンツを融合させる研究を位置づけることができる。テレビと

Webを視聴する端末が共通化され、コンテンツの融合が実現すると、テレビ番組とWebコンテンツを融合してテレビのように視聴したり、逆にブラウザでWebコンテンツに関連付けられたテレビ番組の一コマを閲覧したりと、ユーザが融合されたコンテンツの視聴スタイルを選択できると予想できる(図1参照)。

テレビとWebのコンテンツ融合には、Webをテレビ

番組のように変換 [1][2] したり、テレビ番組を Web コンテンツのように変換 [3] できる技術が活用されることになる。

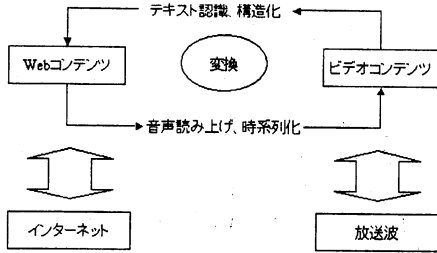


図1 コンテンツと視聴スタイル

コンテンツを変換することによって、対応する端末を持たなかったユーザにも届けられるので、潜在的な視聴者数（リーチ）は伸びる。この場合、元のコンテンツとは違う再生方法になるため、マルチメディア情報の豊かさ（リッチネス）も影響を受ける。しかしながら、コンテンツ変換によってリッチネスやリーチがどう変化するかという研究はあまり行われていない。理論的裏付けを欠いたままでは、コンテンツ変換によりリーチが拡大しても、リッチネスが犠牲にされては視聴者への訴求力に欠ける。そこで本稿では、ブロードバンド化とユビキタス化が同時に進む社会状況を省み、リッチネスとリーチとがバランスを崩さずに発展することを目指したコンテンツ変換を提唱する。

我々は、Web コンテンツをテレビ番組のように視聴するために

- ① テンプレートに基づいて、コンテンツの構造を自動認識し、
 - ② Web コンテンツから文章と画像を抽出分離し、
 - ③ テレビ番組風に提供するための演出を加えて、ビデオコンテンツを生成するシステムを開発している。
- このシステムを用いて実験を行うことで、コンテンツ変換前と変換後で、コンテンツが視聴者に及ぼす影響を比較する。

以下、第2章ではコンテンツ変換の仕組みを解説し、第3章でコンテンツ変換におけるリッチネスを定義する。第4章でリッチネス尺度の計測方法を示し、第5章で知的所有権などサービス実現に臨むための課題を述べる。

2. コンテンツ変換の仕組み

我々は、コンテンツ変換に取り組むにあたり代表的な Web ニュースサイトの HTML コンテンツを自動的にビデオコンテンツに変換する共通プラットフォーム

のプロトタイプを準備した。共通プラットフォームとは、HTML テンプレートを入れ替えることにより、各種の Web サイトに対応できる汎用モジュールの集合である。

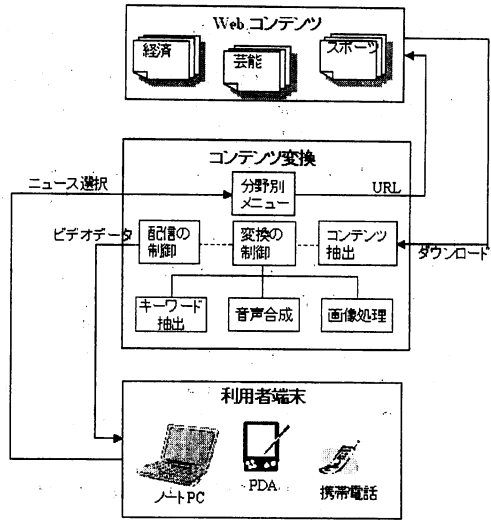


図2 u-PaV システム構成

プラットフォームの中核となる変換サーバは、コンテンツに含まれる画像を、配信先端末の解像度に応じて分割したのち画像処理を行って視聴者に訴えるビデオデータを作り出す。また、コンテンツ中の文章からキーワードを抽出して適切なタイポグラフィを生成する。生成したタイポグラフィは、画像と重ね合わせて強調する。我々はこの共通プラットフォームを u-PaV (ubiquitous/universal Passive Viewer) と称している(図2参照)。

テンプレートに従ってHTMLコンテンツを定型的なビデオコンテンツに変換する u-PaV の他に、テキスト中の語彙を参考に演出方法を選んでコンテンツ変換を行う Web2TV[1] の開発も進めている。Web2TV では、キャラクターアニメーションと音声合成により、Webコンテンツをテレビ番組のようなコンテンツに変換している。

さらに Web2Talkshow[2] では、特定の語彙、あるいは、言い回しに絡めた演出パターンを自動適用し、単調になりがちなビデオコンテンツのなかに視聴者の注意を喚起する対話文を埋め込んでいく。演出には、コンピュータグラフィックスで生成した仮想スタジオを使って行うため、自由度の高い視聴覚刺激を埋め込むことができる。

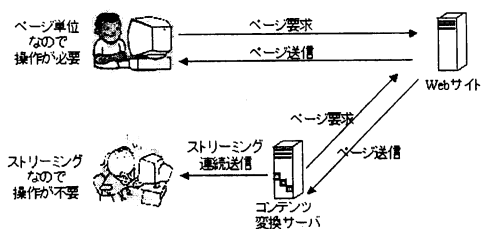


図3 視聴スタイル

こうしたコンテンツ変換サービスを利用できれば、大きなウィンドウで作業しながら、隅のほうに開いた小さなウィンドウで気になるニュース記事をチェックできる。記事ジャンルを選択すると、目次毎にニュースをビデオデータに変換して端末に配信するように自動化すれば、作業をしながら気になるニュースだけ確認できる。

2.1. 視聴スタイル

Web ニュースは主として静止画と文章から構成される。Web ニュースなどの静的コンテンツでは、ユーザがマウスやキーボードを操作して、ブラウザメニューを選択したり、コンテンツに埋め込まれているボタンやハイパーリンクを選択したりする操作によって、先に読み進めることができる。逆の言い方をすると、断続的な働きかけが求められる。つまり、ユーザが使い慣れたブラウザが有効なのは、ユーザの指先が自由になる環境が条件となる。

これに対して、ビデオ映像は、ユーザがチャンネルを選んだら勝手に更新される。Web コンテンツをテレビ風の時系列メディアに変換すれば、働きかけの不要な視聴スタイルになる(図3参照)。

2.2. 変換フロー

コンテンツ変換サーバは、Web サイトからダウンロードしたコンテンツをテキスト/画像/装飾に分離した後、テキストを読み上げた音声と、画像の位置、大きさ、色彩などを変化させた動画を生成し、演出を加えて PC などのユーザ端末にストリーミング配信する(図4参照)。

Web コンテンツはテレビと異なり、動画を含んでいないケースが多いので、ユーザの注意を惹きつけるために演出が求められる。

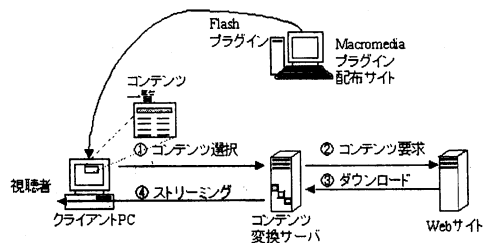


図4 変換フロー

ユーザ端末には、変換されたコンテンツを視聴するためのソフトウェアが必要である。u-PaV の場合には Macromedia Flash 形式に変換したものを視聴するため、ブラウザに Flash プラグインをインストールすることが必要になる。なお、Web2TV の場合は TVML プレイヤー¹[4]が必要になる。

2.3. 変換サービス

情報メディア白書[5]によれば、パソコン以外にもマルチメディアコンテンツを視聴できる端末の種類が増えているが、対価を払ってまでコンテンツを視聴するユーザは伸び悩んでいる。コンテンツの制作コストは高いが、配布コストは安い。コンテンツ総費用のなかで、制作費が大きな固定比率を占め、送信にかかる費用は制作費と比べれば無視できる程度である。

コンテンツ変換は、コンテンツを新規制作するのとは異なり、既存コンテンツを活用するので費用を安く抑えることが期待できる。ビジネスとして成立させるには、価格を抑えて一定数を超える視聴者を確保し、維持することが肝要と考えられる。

コンテンツ費用の回収は、時間が経過するにつれて難しくなるが、とりわけニュース系のコンテンツは、価値の失われるスピードが速い。コンテンツの収益機会を高めるには、情報が新鮮なうちに多くの利用者に販売することが望ましい。コンテンツ変換はこうした方策を後押しするだろう。

変換されたコンテンツは視聴に耐える品質を保つことが求められる。品質が劣化すれば、コンテンツの制作者が伝えようとした情報は、視聴者に正しく伝わらない。コンテンツの情報伝達特性を計るためにリッチネスという考え方を第3章で述べる。

¹ <http://www.nhk.or.jp/strl/tvml/>

3. リッチネスの定義

心理学では、リッチネスとはマルチメディア情報を豊富に含むコミュニケーションを指すとされている。一般に、画像を伴うコンテンツでは、物体の色や形、密度、配置などの情報を認識しやすくなる。認知心理学の分野では、Mandler & Johnson(1976)による“複雑な画像に含まれる情報の種類”の分類[6]、Loftus & Bell(1975)による“画像情報の記憶コード”の分類[7]など比較的古くから研究されてきた。現在では、画像を伴うコンテンツのほうが印象に残りやすい現象は、言語コードと画像コードの両コードを用いて記憶されるという二重符号化説[8]で説明される。音声は記憶に与える影響については、文字情報の単独提示よりも複合提示の場合の方が優れていることを示した研究がある[9]。画像や音声を伴うコンテンツのほうが、文章だけのコンテンツより記憶や認知に有利であるという結論が支持されている。

ただし、過去の研究では、実験者が被験者に写真を直接提示するなど、コンピュータディスプレイを使う現在のマルチメディア環境とは異なる条件で行われており、ユビキタス環境におけるモバイル端末や視聴スタイルの違いまでは考慮されていない。そこで、我々は、ユビキタス環境を反映して、記憶や印象の残りやすさの指標となるリッチネス尺度を検討する。

リッチネスとは、マルチメディア情報が豊富な状態であるには違いないが、リッチネスには他にも様々な要因が入り込む。そこで、はじめにリッチネスの由来を論じたくうえで、端末のマルチメディア表現力を反映させたリッチネス尺度を提案する。

リッチネスの正確な理解を得るために、まず 3.1 節で、リッチネスに纏わる先行研究を紹介する。

3.1. メディアリッチネス

Daft ら[10]は、通信媒体による情報の伝えやすさの違いをメディアリッチネスと呼んだ。

メディアリッチネスは、通信媒体が、以下の4つの要件をどれだけ満たすかで決まる。

- 要件① フィードバックの即時性
- 要件② 映像や音声を伝えられる
- 要件③ 送り手の属性情報の伝達
- 要件④ 言語形式の理解の容易さ

Daft らは、メディアリッチネスの各要件によってもたらされる「情報の伝えやすさ」は、問題解決のために情報を収集する場合の2つの課題を克服するため有効であると唱えた(図5参照)。

1 つめの課題は「不確実性」である。不確実性は、問題を解決するために必要な情報と、現在利用できる情報との差として定義される。

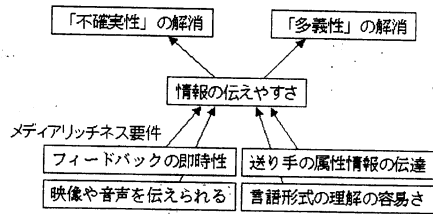


図5 メディアリッチネスの効果

2 つめの課題は「多義性」である。多義性は、ある出来事に対して異なる複数の解釈が存在することを意味する。

メディアリッチネスは、コミュニケーションを対象としており、放送など片方向の伝送形態や Web サイトのコンテンツなど、現在のマルチメディアの発展状況を射程に入れていない。

そこで、コンテンツのリッチネスを論じる場合には、コンテンツの違いを反映できるようにメディアリッチネス要件を修正する必要がある。

3.2 節では、Daft らの主張するメディアリッチネス要件を、コンテンツにどう当て嵌めるかを述べる。

3.2. コンテンツリッチネス

メディアリッチネスの各要件を、以下のようにとらえて、コンテンツにあてはめる。

要件①は、コンテンツのなかの視聴したい箇所が、即座に提示されること(要件①')。

要件②は、コンテンツのなかに映像や音声の情報が含まれていること(要件②')。

要件③は、コンテンツの属性情報がメタデータなどに記述されていること(要件③')。

要件④は、コンテンツ中のテキストに使われている語彙や文法の理解しやすさ(要件④')。

要件①'は、コンテンツの認知に関わる一覧性などを対象とするものである。今回行った実験では、視聴してもらったニュース記事は数百字程度であり、HTML形式のコンテンツを Web ブラウザに表示する場合、記事の任意の位置に視線を移動できるので、視聴方略を自由に選択できる。他方、ビデオコンテンツでは時系列に沿って再生されるコンテンツを視聴することが求められるので方略が制限される。ビデオコンテンツではテキストの表示位置が下方位置に制限される。このように、要件①'は視聴スタイルと密接に関係する。

要件②'は、元のコンテンツに含まれている映像や音声はもとより、コンテンツ変換によって生成される映像や音声も該当する。当然ながら、マルチメディアで伝えるためには、映像や音声を適切に再生できる端末をユーザが持っていることが必須である。

表 1 実験における視聴方式

要件①の 違い	HTML		ビデオ(※)		
要件②の 違い	画像 なし	画像 あり	音声 なし	音声 あり	画像 演出

※ 「画像演出」型の視聴方式は常に音声を伴う。

要件③は、元のコンテンツから画像や文章を正確に抽出するために有益な情報がどれだけ含まれているかを表す。実験では予め指定されたデータ形式で統一しているため、今回の実験では考慮しないことにする。

要件④は、純粋にコンテンツに左右される要件であり、語彙の出現頻度などを調べることで何らかの指標を作成できる。今回の実験では、語彙や文体の難易度が同じ程度になるよう統制し、結果を回答成績の分布で確認する。

4. リッチネス尺度

リッチネス尺度は情報の伝えやすさであるから、コンテンツに関する設問の回答成績に反映されると考える。今回の実験では、リッチネス尺度を決定する要因としてコンテンツのリッチネス要件①'と要件②'を考慮する。すなわち、要件①'を x_1 、要件②'を x_2 において $y = f(x_1, x_2)$ 型の関数でリッチネス尺度を検討する。

4.1. 実験計画

我々は、まず要件①'の違いに対応して、ニュース記事全文が Web ブラウザ 1 画面の幅・高さ収まる HTML レイアウトによるコンテンツ視聴方式と、時系列再生されるビデオによる視聴方式の 2 系統に分類した。

次に要件②'の違いに対応して、HTML による視聴では、画像を含むケースと含まないケースの 2 種類を用意した。ビデオによる提示方法では、映像を共通にした音声ありと音声なし、そして、音声と映像演出を加えた 3 種類を用意した (表 1 参照)。

架空のニュース記事を 3 本用意し、視聴方式ごとにコンテンツを変換して提示する。視聴方式の違いがニュース記事からの情報獲得に如何なる影響を与えているかを評価する。

リッチネスは、情報の多義性や不確実性を減らし、情報を有効に伝達すると考えられる。そこで被験者がコンテンツを視聴した後に、理解と記憶に対する効果を調べる設問を被験者に課す。

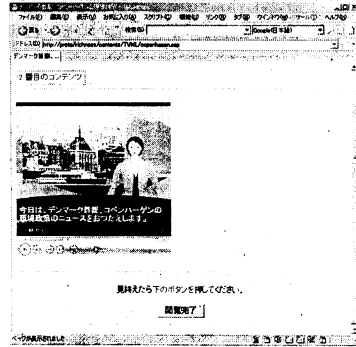


図 6 Web2TV 視聴方式

今回、実験に参加してもらったのは、10 代後半から 20 代前半の学生 (男 : 女 = 2 : 1 前後)、ならびに、20 代から 40 代の事務職の女性である。

ただし、理解力や記憶力には個人差がある。そこで、視聴方式を揃えたコンテンツを用意して、被験者の能力差を共通問題の成績に基づき補正する。また、態度によって記憶に差が生じることを鑑み、アンケートによってコンテンツへの興味等を事後評価する。即ち、コンテンツを視聴してもらった後に

「あなたの興味を引く情報が含まれていましたか」
 「あなたにとって役立つ情報が含まれていましたか」
 「あなたにとって新しい情報が含まれていましたか」というアンケートに答えてもらう。これらの項目群を順に、興味度、有用度、新規度と呼ぶことにする。

ここで、興味度と有用度は被験者の視聴態度に正の効果を与え、新規度は視聴態度に負の効果を与えると考えられる。なぜなら、視聴者にとって既知の内容は視聴意欲を下げるはずである。

4.2. 評価システム

被験者には、インターネット経由で実験用サイトに被験者ごとに異なる ID でログインしてもらい、変換後のコンテンツを、所定の視聴方式でアクセスさせる (図 6 参照)。

その直後に、視聴結果に対して興味度、有用度、新規度のアンケートを行う。アンケートに続いて理解力や記憶力を調べる設問が提示される。設問は、選択式と記述式の両方を用意した。結果の中間集計によると、選択式の設問ではいずれの視聴方式でも成績が良好で差があまり認められないのに対して、記述式の設問では、画像・音声・演出を伴うコンテンツのほうが良好な成績となっている (表 2, 表 3 参照)。

表 2 選択式設問の回答成績

	HTML		ビデオ		
	画像なし	画像あり	音声なし	音声あり	画像演出
正解	12	10	12	19	13
不正解	1	1	3	1	3

表 3 記述式設問の回答成績

	HTML		ビデオ		
	画像なし	画像あり	音声なし	音声あり	画像演出
正解	4	7	8	12	10
不正解	9	4	7	8	6

5. まとめ

コンテンツの理解しやすさを定量化する試みは、心理学の分野で様々な実験が行われてきた。心理学実験の厳密性はおろそかにすべきではないが、我々が日常新しい情報を吸収するのに用いているテレビや Web の視聴スタイルを評価する尺度として当て嵌めるのは困難である。ユーザが理解しやすいコンテンツを探す目安として、我々が提案するリッチネス尺度のようなものがあれば便利である。

元のコンテンツ形式に対応していない端末でも、コンテンツを変換すれば視聴可能となるが、変換の精度や演出によってマルチメディア情報の豊かさは変化し、それをどの程度楽しめるかは端末仕様の制限も受ける。依然、携帯電話の機種によって動画や音声の再生に制限があるので、リッチネス尺度のような簡便な指標でも役に立つだろう。

ところで、コンテンツ変換サービスを実現するには著作権法で規定される公衆送信権や送信可能化権を侵害しないことを保証しなければならない[11]。コンテンツの変換出力を、変換を求めた者だけが利用可能であることを保証し、私的使用と認められなくてはならない。著作者の承諾を得なくても著作物を利用できる行為として、著作権法第 30 条に「私的使用のための複製」がある。変換後のコンテンツが、この類型とあると認めらる必要がある。

パソコンで誰でも視聴できるコンテンツを、視聴スタイルや端末のタイプを超えて視聴できる状態にすることが誰かの利益を侵害する懸念は小さい。まず、変換したコンテンツを第三者に再送信するのではないから、著作権者の利益とは相反しない。尤も広告を除かれると困る人がいるかもしれないが、ニュースと広告は元来別の内容なので、それらを分離することは著作者人格権の同一性保持(第 20 条)に抵触しないだろう。ただし、コンテンツ変換に伴い演出を加えることは、

著作者人格権に抵触する恐れがある。こうした議論を深めてサービス実現に向けた課題をねばり強く解決していく心構えが求められる。

参考文献

- [1] 灘本 明代, 服部 多栄子, 近藤 宏行, 沢中 郁夫, 田中 克己, “Web コンテンツの受動的視聴のための自動変換とスクリプト作成マークアップ言語”, 情報処理学会論文誌: データベース (TOD8) Vol.42, No.SIG1, pp.103-116 (2001)
- [2] 蓬萊 博哉, 灘本 明代, 田中 克己, “理解しやすさとユーモアを考慮した Web コンテンツの対話文変換”, 日本データベース学会 Letters Vol.2, No.2, pp.29-32 (2003)
- [3] 宮森 恒, 田中 克己, “ウェブ化ビデオ: テレビ番組からウェブコンテンツへのメディア変換および関連情報との統合的閲覧方式”, データ工学ワークショップ(DEWS2005), 2C-i6, (2005)
- [4] 林 正樹, “番組記述言語 TVML を使った情報の番組化”, 情報処理学会 DBS 研究会技術報告, 2000-DBS-120-13, Vol.2000, No.10, pp.91-98, (2000)
- [5] 電通総研, “情報メディア白書 2004”, (2004)
- [6] Mandler, J. et Johnson, N. Some of the thousand words a picture is worth. Journal of Experimental Psychology: Human Learning and Memory, 2, pp.529-540 (1976)
- [7] Loftus, G.R. and Bell, S.M., “Two types of information in picture memory”, Journal of Experimental Psychology: Human Learning and Memory, 1, pp103-113 (1975)
- [8] Paivio, A., “Imagery and synchronic thinking”, Canadian Psychological Review, 16, pp147-163 (1975)
- [9] 五十島愛理, 筒井雄二, “聴覚刺激と視覚刺激を用いた短期記憶課題におけるモダリティ効果 再認法を用いた分析”, 日本心理学会大会発表論文集, Vol.65th, JN:X0238A, (2001)
- [10] Daft, R.L. & Lengel, R.H., “Information richness: a new approach to managerial behavior and organizational design”, Research in organizational behavior 6, pp191-233 (1984)
- [11] 大阪弁護士会 知的財産法実務研究会 編著, “デジタルコンテンツ法 [上巻]”, pp129-145 (2004)