

マクロブロックのレート・歪み特性に着目した 高効率 2 パス符号化アルゴリズムの検討

横 張 和 也^{†1} 嵯 峨 田 淳^{†1} 上 倉 一 人^{†1}

高圧縮符号化を実現するために、2パス符号化が提案されている。既存の2パス符号化では、符号化結果のフレームごとの特徴量に着目することでフレーム単位の符号化効率を向上させ、高圧縮を実現している。本研究では、マクロブロックの符号化効率（レート歪特性）に着目する。1パス目の符号化結果をもとに、各マクロブロックの符号化効率を最適化する量子化パラメータを決定し、2パス目の符号化に反映することで、マクロブロック単位での符号化効率を向上させ、高圧縮を実現するアルゴリズムを提案する。提案手法を用いることで1パス符号化の結果と比べ平均約20%の符号化効率向上を確認した。

High-Efficient 2-pass Video Coding Algorithm based on Macroblock Rate-Distortion

KAZUYA YOKOHARI,^{†1} ATSUSHI SAGATA^{†1}
and KAZUTO KAMIKURA^{†1}

In order to achieve high-efficient video coding, 2-pass video coding algorithms are proposed. In the existing 2-pass video coding algorithms, To focus on the feature quantity of the coding result in the frame, the video coding efficiency is grown in the frame. In this paper, we focus on the macroblock coding efficiency (rate-distortion). In the proposed algorithm to achieve high-efficient video coding, the quantization parameters which optimize video coding efficiency of each macroblocks are decided based on the 1-pass encoding result. Using these quantization parameters at the 2-pass encoding, the video coding efficiency is grown in the macroblock. The proposed algorithm is able to improve the video coding efficiency by about 20 percents.

1. はじめに

現在、映像コンテンツの利用シーンは多岐に渡っており、その空間解像度も様々である。携帯電話では QVGA (320 × 240) や WVGA (800 × 480)、地上波デジタル放送・BS デジタル放送では 720 × 480, 1440 × 1080, 1920 × 1080 などの空間解像度が用いられている。それぞれのアプリケーションにおいて、ユーザーから求められている空間解像度は拡大していく傾向がある。その一方で、映像コンテンツの伝送帯域にはアプリケーションに応じた制限があることが多いため、映像符号化方式に対しては、より高圧縮を実現することが真に望まれている。

高圧縮を実現する符号化アルゴリズムとしては、2パス符号化アルゴリズムが知られている¹⁾。2パス符号化アルゴリズムでは、第一符号化器が時間方向に先行して解析を行い、符号量を再配分した上で、第二符号化器が再符号化することで高圧縮を実現している²⁾。

従来の2パス符号化アルゴリズムでは、第一符号化の結果を時間方向に解析することで未来のフレームの複雑さ指数を取得したり、シーンチェンジの検出を行うことで高圧縮化を実現している。その為、解析対象となるフレーム数が多くなるほど圧縮率向上の効果が大きくなる。一方で、従来より提案されている方式ではフレーム単位での解析・符号量再配分を行い高圧縮化を実現しており、フレーム内のマクロブロック単位での符号量最適化は行われていない。

そこで、マクロブロック単位で符号量再配分を行い、更なる高圧縮化を実現する2パス符号化アルゴリズムを提案する。提案手法では、符号量再配分を行うにあたり第一符号化結果による符号化済映像に対して客観評価を行う。第一符号化結果と客観評価より得られた評価値を元に、第二符号化での量子化パラメータを決定し、その量子化パラメータを用いて第二符号化を行うことで高圧縮符号化を実現する。

以降、2章にて提案方式の概要について述べる。3章にて提案方式のアルゴリズムについて述べる。4章にて提案方式の実験および考察を述べ、5章において本稿をまとめる。

2. 提案手法のアプローチ

従来の2パス符号化では、フレーム単位での解析・符号量再配分を行うことにより高圧

^{†1} 日本電信電話株式会社 NTT サイバースペース研究所
NTT Cyber Space Laboratories, NTT Corporation

縮化を実現していたため、マクロブロック単位での符号量最適化は行えていなかった。そこで、マクロブロック単位での符号量最適化を行い、更なる高圧縮化を実現する新たな2パス符号化アルゴリズムを検討した。

ある符号化対象ピクチャを符号化するにあたり、ピクチャの総発生符号量対総歪みの観点で符号化効率が最適となるのは、Lagrangeの未定乗数法から明らかなように各マクロブロックの符号化効率が等しい時、すなわち

$$\frac{\partial D_{mb}}{\partial R_{mb}} = \frac{\partial D_{picture}}{\partial R_{picture}} = constant \quad (1)$$

を満たす時である³⁾。これは各マクロブロックの発生符号量に対する画質への貢献を均等化することを意味する。

提案するアルゴリズムでは符号化効率を最適化するために、全てのマクロブロックにおいて(1)式を満たすように各マクロブロックの量子化パラメータの調整を行い、マクロブロック単位の発生符号量を最適化することを目指す。各マクロブロックで(1)式を満たすためには、各マクロブロックのレート歪み特性(RD特性曲線)を知る必要がある。しかし、RD特性曲線について定式化されたものは存在していない。そこで、本検討ではまず第一符号化結果を元にRD特性曲線を近似する方式について検討する。また、第二符号化時において目標とする符号化効率($\partial D/\partial R$)を満たす量子化パラメータを導出するために、同様にQPと符号量の関係を表すQP-GenBit曲線の近似式についても事前に検討した。

2.1 客観評価手法：SSIM

提案手法では、第一符号化結果の評価を行う際に客観評価を用いる。今回は客観評価手法としてSSIM⁴⁾を用いた。SSIMは従来から客観評価手法として一般的に使用されているPSNRと比べて、より主観評価に近い結果が得られるとされている手法である。PSNR⁵⁾は客観評価手法としてよく用いられてきたが、近年、SSIMがAVC/H.264リファレンスソフトであるJM⁶⁾やGNU GPL配布のAVC/H.264ソフトエンコーダであるx264²⁾などでも実装されてきており、一般的に用いられつつある。

SSIMは以下の様な式で定義される。ここで、 x, y はそれぞれ画像を表しており、 x, y はそれぞれのマクロブロックの輝度値を表している。

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (2)$$

式を構成する要素はそれぞれ以下の通り定義される。

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (3)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (4)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x^2\sigma_y^2 + C_3} \quad (5)$$

ここで、 $l(x, y)$ は輝度値比較関数、 $c(x, y)$ はコントラスト値比較関数、 $s(x, y)$ 構造比較関数と呼ばれている。また、ここで用いられている $\mu_x, \mu_y, \sigma_x, \sigma_y, \sigma_{xy}$ はそれぞれ、以下の通り定義される。

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (6)$$

$$\mu_y = \frac{1}{N} \sum_{i=1}^N y_i \quad (7)$$

$$\sigma_x = \left\{ \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right\}^{\frac{1}{2}} \quad (8)$$

$$\sigma_y = \left\{ \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2 \right\}^{\frac{1}{2}} \quad (9)$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (10)$$

ここで、 μ_x, μ_y は対象領域内の輝度値の相加平均、 σ_x, σ_y は対象領域内の輝度値の分散、 σ_{xy} は対象領域内の輝度値の共分散を表している。また、 $C_1, C_2, C_3, \alpha, \beta, \gamma$ はそれぞれ補正項を表しており、 $C_1 = 6.5025, C_2 = 58.5225, C_3 = C_2/2, \alpha = \beta = \gamma = 1$ と定義される⁴⁾。

2.2 RD特性近似曲線

第一符号化結果から得られた各マクロブロックの発生符号量および客観評価より得られた各マクロブロックのSSIMから、第二符号化を行う際にターゲットとするマクロブロック発生符号量を求めるために、マクロブロックの発生符号量とSSIMについてのRD特性近似曲線を定義する。

RD特性近似曲線を定義するため、様々なマクロブロックについて、QPを0から51まで設定した時の発生符号量とSSIMの関係をグラフにプロットし、実際のRD特性を確認

した．結果を図 1 に示す．プロットの形状から $y = x^n$ のグラフとの相関が高いことがわかる．そこで，発生符号量と SSIM の関係を表す RD 特性近似曲線を以下の (11) 式のように近似する．

$$f(x) = 1 - (1 - x)^a \quad (11)$$

ここで， x はマクロブロック最大発生符号量 (3200bit) で正規化した発生符号量， $f(x)$ は SSIM が入り， a はマクロブロックごとに異なる値をとる変数である．任意の QP でマクロブロックを符号化した結果得られる発生符号量と SSIM を (11) 式へ与えることで a を決定し当該マクロブロックの RD 特性近似曲線を求めることができる．

上で求めた RD 特性近似曲線 $f(x)$ の精度を検証した．検証方法は以下の通りである．あるマクロブロックを 1 つの代表 QP 値で符号化した時の発生符号量・SSIM を抽出し， $(x_k, f(x_k)) = (Mbit, SSIM)$ を代入して a_k を求める．次に，求められた当該マクロブロックの RD 特性近似曲線 $f(x) = 1 - (1 - x)^{a_k}$ の x に QP が 0 から 51 までの時の実発生符号量を代入し， $f(x)$ (= 近似 SSIM) を得る．得られた近似 SSIM と，実際に QP を 0 から 51 までの値の時に得られる SSIM について相関係数を求めた．

あるマクロブロックについて実際の RD 特性曲線と RD 特性近似曲線をプロットしたものを図 3 に示す．RD 特性曲線と RD 特性近似曲線を比べると，同様の傾向を持つ曲線となっていることが確認出来る．

代表点として 3 つのマクロブロックを選定し，相関係数を求めたところ最大で 0.994，最小で 0.975 という結果が得られた．いずれも 1 に非常に近い値が相関係数として得られているため， $f(x)$ は十分 RD 特性曲線の近似式として適用出来ると考えられる．

2.3 QP-発生符号量曲線 (QP-Bit 曲線)

先に求めた RD 特性近似曲線を用いて決定された第二符号化時の各マクロブロックのターゲット符号量となるような QP を求めるために，QP と発生符号量の関係を表す QP-発生符号量曲線 (QP-Bit 曲線) の近似式について検討した．

RD 特性近似曲線を定義したときと同様に，様々なマクロブロックについて，QP を 0 から 51 まで設定した時の，QP と発生符号量の関係をグラフにプロットし，実際の QP-発生符号量の関係を確認する．結果を図 2 に示す．プロットの形状からそれぞれの曲線がほとんど交差することがないことがわかる．そこで，代表的な点を用いて QP と発生符号量の関係を表す QP-Bit 曲線を以下の 2 次方程式で近似する．

$$g(x) = \alpha x^2 + \beta x + \gamma \quad (12)$$

ここで， x は QP， $g(x)$ はマクロブロック発生符号量が入る．今回は代表的な点として様々

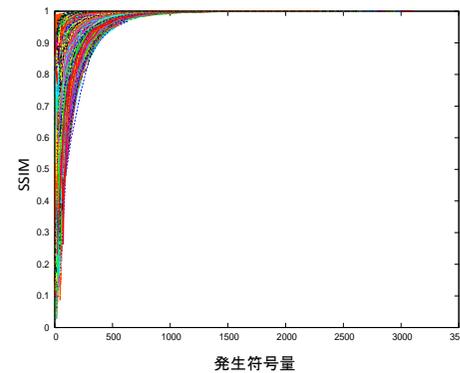


図 1 RD 特性曲線
Fig. 1 RD Curve

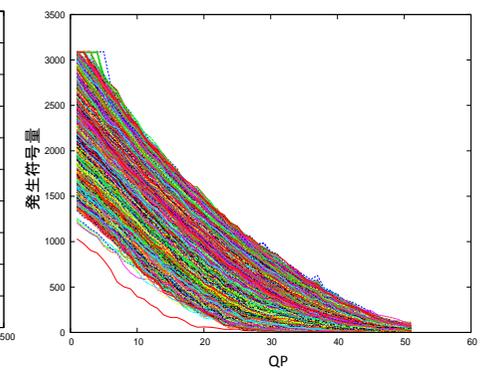


図 2 QP と発生符号量
Fig. 2 QP-Generated Bits

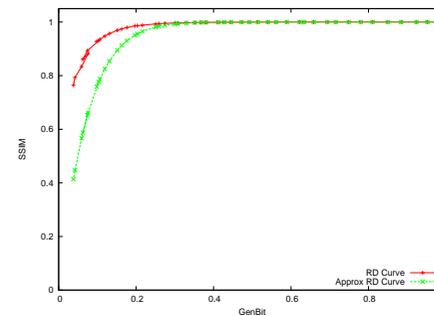


図 3 RD 特性近似曲線
Fig. 3 Approx RD Curve

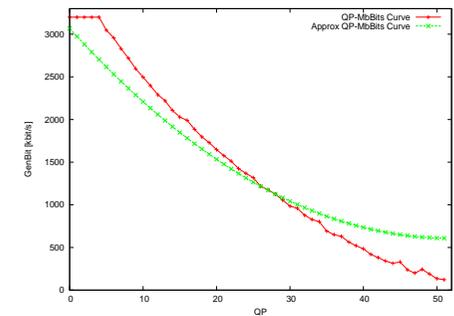


図 4 QP-発生符号量曲線
Fig. 4 QP-Generated Bit Curve

な QP-GenBit 曲線の平均値を求め，その平均値と (12) 式に対して最小自乗法を用いて $\alpha = 0.91846975, \beta = -95.079875$ と求めた．また， γ は y 軸切片を表しており曲線への依存性が強い値であると考えられるため， γ を変数として任意の QP とマクロブロック発生符号量を (12) 式へ与えることで 1 本の QP-GenBit 曲線を決定する．すなわち，第一符号化結果を与えることで各マクロブロックの QP-GenBit 曲線を決定する．

RD 特性近似曲線の時と同様に，上で求めた QP-Bit 近似曲線 $g(x)$ の精度を検証した．検証方法は以下の通りである．あるマクロブロックを一つの代表 QP 値で符号化した時の発生符号量 (MbBit) を抽出し， $(x_i, g(x_i)) = (QP, MbBit)$ を代入して， γ_l を求める．次に，求められた当該マクロブロックの QP-MbBits 近似曲線 $g(x) = \alpha x^2 + \beta x + \gamma_l$ の x に QP 値 (= 0 から 51) を代入し，その時に得られる近似発生符号量と，実際の QP 値の時の発生符号量について相関係数を求めた．

あるマクロブロックについて実際の QP-Bit 曲線と QP-Bit 近似曲線をプロットしたものを図 4 に示す．QP-Bit 曲線と QP-Bit 近似曲線を比べると，同様の傾向を持つ曲線となっていることが確認出来る．

代表点として 3 つのマクロブロックを選定し，相関係数を求めたところ最大で 0.999，最小で 0.993 という結果が得られた．QP-Bit 近似曲線も RD 近似曲線と同様に，いずれのもの 1 と非常に近い値が相関係数として得られているため， $g(x)$ は十分 QP-Bit 曲線の近似式として適用出来ると考えられる．

3. アルゴリズム説明

提案するアルゴリズムでは，第二符号化のターゲット発生符号量およびターゲット SSIM を決定し，RD 特性近似曲線上でターゲットを通る点の傾き $\partial D / \partial R$ を求める．求めた $\partial D / \partial R$ を全てのマクロブロックが満たすように符号量を調整し，その符号量に対して QP-Bit 曲線から QP を求める．求めた QP を用いて第二符号化を行うことで，全てのマクロブロックの符号化効率の最適化を実現する．

提案するアルゴリズムのフローチャートを図 5 に示す．図 5 を用いて詳細な動作について説明する．

1. 第一符号化 入力画像に対して符号化を行い符号化ストリームを出力する．
2. 復号 第一符号化結果のストリームの復号を行う．復号時には，復号映像信号の出力と共に，第一符号化時のマクロブロックタイプおよびマクロブロック発生符号量を符号化情報として出力する．
3. 評価 復号結果より得られた復号映像信号と入力映像信号に対して客観評価を用いて，ピクチャおよびマクロブロックごとの客観評価値を求める．今回は，客観評価手法として SSIM を用いるので，ピクチャおよびマクロブロックごとの SSIM を求める．
4. 目標符号化効率の算出 各マクロブロックのターゲット発生符号量を求める前に，任意の第二符号化時のターゲット符号化効率 $E_{pic} = \partial D_{pic} / \partial R_{pic}$ を定める．

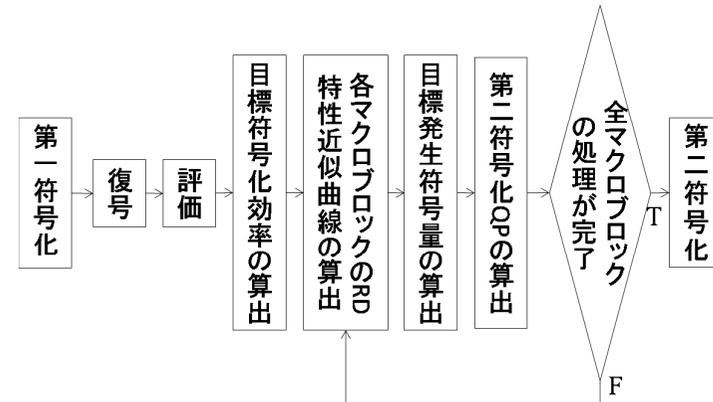


図 5 フローチャート
Fig. 5 Flow chart

後述する本稿の実験において， E_{pic} は第一符号化時の平均ピクチャ発生符号量 (AvePicBit) と平均 SSIM (AveSSIM) の値から導出した．AvePicBit および AveSSIM を (11) 式より導かれた符号化効率近似式 ((13) 式) へ代入し，以下のようにターゲットとなる符号化効率 E_{pic} を求める．

$$E(x) = f'(x) = a(1-x)^{a-1} \quad (13)$$

$$E_{pic} = a_{pic}(1 - AvePicBit)^{a_{pic}-1}, \text{ here} \quad (14)$$

$$a_{pic} = \frac{\ln(1 - f(x))}{\ln(1 - x)} = \frac{\ln(1 - AveSSIM)}{\ln(1 - AvePicbit)}$$

5. 各マクロブロックの RD 特性近似曲線の算出 第一符号化時のマクロブロック発生符号量 (x_i) と，評価時に得られた SSIM を (11) 式へ代入し，マクロブロック固有の変数 a_i を求め，マクロブロック RD 特性近似曲線を求める．

$$f(x_i) = 1 - (1 - x_i)^{a_i-1}, \text{ here } a_i = \frac{\ln(1 - f(x_i))}{\ln(1 - x_i)} \quad (15)$$

- ここで i は i 番目のマクロブロックを意味する添え字である．
6. 目標発生符号量の算出 5. で求めた a_i によって定められた各マクロブロックの RD 特性近似曲線上において，符号化効率が 4. で定めた E_{pic} となる発生符号量を求める．ここで求めた発生符号量が各マクロブロックの第二符号化時の目標とする符号量 ($TMbBit_i$) となる．(13) 式へ $E(x_i) = E_{pic}$ ， $a = a_i$ を代入し $TMbBit_i$ を求める．

$$\begin{aligned}
 E(x_i) &= a_i(1-x_i)^{a_i-1} \\
 E_{pic} &= a_i(1-TMbBit_i)^{a_i-1} \\
 TMbBit_i &= 1 - \left(\frac{E_{pic}}{a_i}\right)^{\frac{1}{a_i-1}}
 \end{aligned} \tag{16}$$

7. 第二符号化 QP の算出 (11) 式で求めた QP-GenBit 近似式へ、復号時に得られたマクロブロックの QP (QP_i) と発生符号量 ($MbBit_i$) を代入し、 γ_i を求め、QP-GenBit 近似曲線を求める。

$$\gamma_i = MbBit_i - \alpha \cdot QP_i^2 - \beta \cdot QP_i \tag{17}$$

求めた QP-GenBit 近似曲線へ γ_i を代入した後、 $g(x_i) = TMbBit_i$ を代入して解の公式より第二符号化時に用いる QP (TQP_i) を求める。このとき、 TQP_i の値域は $0 \leq QP \leq 51$ と定義されるため、値を一意に求めることができる。

8. 第二符号化 第一符号化結果より得られた符号化情報に対して、7. で求めた第二符号化時の QP (TQP_i) を適用し第二符号化を行い、符号化ストリームを出力する。

4. 実験

3章で述べた提案手法の有効性を確認するために、12種類の ITE 標準画像を用いて実験を行った。画像サイズは SDTV サイズ (720 × 480)、フレーム数は 450 フレーム、ピクチャ構造は I-ピクチャのみで実験を行った。また、実験では JM の ver. 14.2 を使用した。

本実験では、3章で述べた通り $E_{pic} = \partial D_{pic} / \partial R_{pic}$ を決定する際に、対象とするフレーム内の全マクロブロックの平均発生符号量 (AvePicBit) および、平均 SSIM (AveSSIM) を用いている。第一符号化と第二符号化の結果に対して、RD 特性曲線を取得した。実験結果の一つ (Soccer Action) を図 6 に示す。Soccer Action では、第二符号化時での QP の値は第一符号化時と比べて最大で ±28 の変化域があり、符号化効率は全体で約 23% 向上した。

5. まとめ

本稿では、マクロブロックごとに符号量の再配分を行い、符号化効率を向上させる 2パスアルゴリズムを提案した。提案したアルゴリズムでは符号量の再配分を行う際、客観評価手法の一つである SSIM を用いた。本アルゴリズムを適用することで、最大で約 20% 符号化効率が向上することを確認した。一方で、実験で用いた 12 種類の画像の中で効果があまり得られなかった画像も存在していたため、今後は更なる符号化効率の向上を目指し、符号量再配分アルゴリズムの改良および、インターへの解析結果の利用を検討していく。

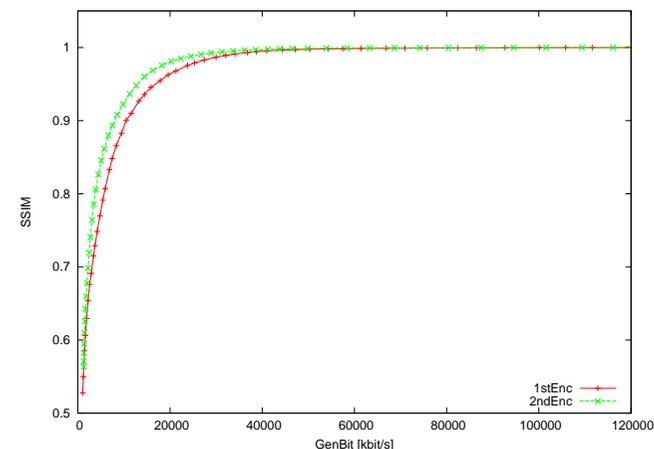


図 6 実験結果 (Soccer Action)

Fig.6 Experimental Result (Soccer Action)

参考文献

- 1) Yutaka Yokoyama, Satoshi Nogaki, "A Rate Control Method with Pre-Analysis for Real-Time MPEG-2 Video Coding," *IEEE CNF*, vol. 3, pp. 514-517, 7-10 Oct. 2001.
- 2) H.264/AVC 用符号化ライブラリ, x264, "<http://www.videolan.org/developers/x264.html>" (2010), 閲覧可能
- 3) Seishi Takamura, Naoki Kobayashi, "MPEG-2 one-pass variable bit rate control algorithm and its LSI Implementation," *Image Processing, 2001. Proceedings. 2001 Internatioanl Conf.*, vol. 1, pp. 942-945, 7-10 Oct. 2001.
- 4) Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- 5) Huynh-Thu Q, Ghanbari M, "Scope of Validity of PSNR in Image/Video Quality Assessment," *IEEE Electronics Letters*, Vol. 44, Issue. 13, pp. 800-801, June 19 2008.
- 6) Joint Model Reference Software, JM, "<http://iphome.hhi.de/suehring/tml/>" (2010), 閲覧可能