

人工的な多値画像のベクタ変換を考慮した領域分割に関する検討

河村 圭[†] 石井 大祐[†] 渡辺 裕[†]

[†] 早稲田大学大学院 国際情報通信研究科
〒 367-0035 埼玉県本庄市西富田 1011 A310
E-mail: †kawamura@suou.waseda.jp

あらまし オブジェクト符号化は、定常状態の条件を満たす領域群に画像を分割して符号化する。この符号化の境界記述方式としてベクタ表現の適用が考えられる。一方、我々はこれまでに解像度変換が容易なベクタ表現に着目し、小数画素精度の2値化による境界線抽出手法とそのベクタ変換を提案している。対象画像は、本来2値画像であるが、画像取得時に低解像度多値画像として得られた画像である。本稿では、本来均等色領域から構成されていた多値画像が、画像入力時にアンチエイリアシングを含む多値画像（人工画像）として得られた画像を対象とする。そして、ベクタ表現による画像符号化を目的として、人工画像の特徴を考慮した領域分割手法を提案する。まず、アンチエイリアシングを考慮した領域分割を実現するためのノルムを導入する。次に、提案したノルムを用いた目的関数の最小化問題へと帰着させ、これを近似的に解くアルゴリズムを示す。最後に、近年注目されているTVノルムによる骨格画像との比較実験を行い、その結果と考察を述べる。

キーワード 領域分割, 均等色領域, アンチエイリアシング, ベクタ表現, 画像符号化。

A Study on Segmentation of Artificial Grayscale Image for Vector Conversion

Kei KAWAMURA[†], Daisuke ISHII[†], and Hiroshi WATANABE[†]

[†] Graduate School of Global Information and Telecommunication Studies, Waseda University,
A310, 1011 Nishi-Tomida, Honjo-shi, Saitama 367-0035, Japan.
E-mail: †kawamura@suou.waseda.jp

Abstract Object oriented coding segment an image into regions, which characteristics are coherent. Vector representation is suitable to describe contours of regions in object oriented coding. On the other hand, we have proposed border extraction using sub-pixel binarization for the low resolution grayscale image. This image which is originally binary image is obtained as grayscale image by a scanner. In this paper, a target is an image which originally consists of homogeneous regions is obtained as grayscale image containing anti-aliasing. This image is named the “artificial image.” The research object is the image coding based on vector representation. We propose the segmentation method considering characteristics of the artificial image. An “inverse norm” is introduced to achieve the segmentation. Then the segmentation problem is converted to the minimization problem of an object function with the inverse norm. An algorithm to solve the minimization problem is designed. Total variation image decomposition is recently remarkable technique as preprocess of segmentation. The validity of the proposed method is confirmed by comparative experiments, which are proposed method and total variation.

Key words Image segmentation, homogeneous region, anti-aliasing, vector representation, image coding.

1. はじめに

現在、大部分の静止画像符号化は変換符号化（と予測符号化）を基本としている。国際標準規格である JPEG では 8×8 のブロックごとに DCT (Discrete Cosine Transform) が、JPEG2000 では画像全体に DWT (Discrete Wavelet Transform) が適用される。また、H.264 のイントラフレームでは 4×4 のブロック単位にイントラ予測と整数化された DCT が適用される。これらの方式では画像に含まれる性質が異なる領域の境界を十分考慮しないことで、計算複雑度を低減させている。特に、固定ブロックサイズを用いると、以下の二つの問題がある。まず、それぞれのブロック内が定常状態であるとは限らないため、符号化効率が低下する。さらに、ブロックを独立して扱うため、復号画像でブロック境界が知覚されてしまう。なお、画像全体を用いても、前者の問題は依然として生じる。従って、さらなる符号化効率の向上が可能であると思われる [1]。

無理な定常状態を仮定している問題点を解決するための方式として、オブジェクト符号化が提案されている [2]。この方式では、画像を与えられた定常状態の条件を満たす領域群に分割して符号化する。分割されたそれぞれの領域を独立して扱うため誤分割が生じると、その境界が強調されてしまう。高い主観品質を実現するためには、視覚的な分割結果に出来るだけ近い領域分割手法を適用する必要がある。古くは SIC (Segmented Image Coding) として、境界情報の記述方式や領域分割に関して多くの研究がなされ [2]、MPEG-4 Part 2 の一部として国際標準化されている。近年では MRC (Mixed Raster Content) として、混在画像に対する多くの研究がなされ [3]、JPEG2000 Part 6 として国際標準化されている。

オブジェクト符号化における境界記述方式としてベクタ表現の適用が考えられる。我々はこれまでに解像度変換が容易なベクター表現に着目し、2 値化による境界線抽出及び境界線のベクター変換手法を提案している [4-7]。その背景として、解像度の異なる様々な表示端末の増大がある。これらの端末にコンテンツを表示するには解像度変換が必須である。また、コンテンツとして写真などの自然画像だけでなく、文書画像や CG、アニメーションなどの人工的な画像へと広がっている [8]。

画像の取得時においてスキャン時間は短い方が良いため、低解像度が用いられることが多い。さらに JPEG 符号化方式の普及により、文書画像が低解像度多値画像で保存される機会が増加している。そこで我々は本来 2 値画像であるが、画像取得時（デジタル化した時）に低

解像度多値画像として得られた画像に対して、小数画素精度の 2 値化を適用する手法を提案してきた [7]。2 値化による境界線抽出は、MRC に適用されている領域分割の概念に近い。

さらに対象画像を広げるために、本来 2 値であるという制限を緩和する。すなわち、本来均等色領域から構成されていた多値画像が、画像入力時にアンチエイリアシングを含む多値画像として得られた場合を想定する。以下では、このような画像を自然画像に対して人工画像と呼ぶ。人工画像はアニメーションやイラストなどに見られる。また、自然画像が重畳されている場合もある。2 値化に限定されない境界線抽出は SIC に適用されている領域分割の概念に近い。

本稿では SIC の概念に基づき、ベクトル表現による画像符号化のための人工画像を考慮した領域分割手法を提案する。まず、符号化に関連した領域分割手法のうち MRC と SIC について述べる。さらに、近年注目されている TV ノルムによる骨格画像について述べる。次に、アンチエイリアシングを考慮した領域分割を実現するためのノルムを提案する。また、領域分割を提案したノルムを用いた目的関数の最小化問題へと帰着させ、これを近似的に解くアルゴリズムを示す。最後に、実験結果と考察を述べる。

2. 関連研究

2.1 MRC

MRC (Mixed Raster Content) は静止画像を前景、背景、マスクの 3 レイヤーに分離する概念である。特に文書 (2 値画像) と写真 (階調画像) が混在している混在画像 (Compound Image) を対象としている。また、各レイヤーで解像度と符号化方式を最適化することで、高い符号化効率を可能とする。本手法の課題は、文書画像に相当するマスクレイヤーを効率よく生成すること、前景と背景レイヤーの充填問題である。本稿では特に前者を取り上げる。

Ricardo らはマスクを生成するために、レート歪み特性に基づいて 2 値化の最適なしきい値を求める手法を提案している [3]。画像を 8×8 画素のブロック $\{x_n(i, j)\}$ に分割し、最適なマスク $\{m_n(i, j)\}$ を探索する。それぞれのブロックについて 2^{64} 通りのマスクが考えられる。レート歪みの観点から最適なマスクは、コスト関数

$$J_n = R_n + \lambda D_n \quad (1)$$

を最小化することになる。 λ は与えられたブロックについてレート歪みのトレードオフを制御する係数である。

ここで、文書画像の最も単純なモデルとして画素のヒストグラムがある。背景とテキストはヒストグラムで二

つの峰を持つはずであるが、双峰性検定の手法にかかわらず、画素は何らかのしきい値によって分類される。ブロック内でしきい値処理によりマスクを得ることは、

$$m_n(i, j) = u(t_n - x_n(i, j) - 1) \quad (2)$$

で表される。ここで、 t_n はブロックのしきい値であり、 $u(k)$ は離散ステップ関数 ($k \leq 0$ の時に 1, それ以外は 0) である。結果として、しきい値以下の画素は背景レイヤーに分離される。ブロック内には 64 個の画素があるため、64 個の有意なしきい値が存在する。しきい値処理により得られるマスク群は、 2^{64} 通りマスクのうち最適なマスクを含んでいることが既に示されている [9]。

本手法は、画像のモデル化において前景と背景の 2 種類しか仮定されていない。また前景の個数を拡張可能であるが、ブロック間でのしきい値の連続性が保証できない。したがって、MRC を複数の均等色を含む人工的な画像に適用するのは困難である。さらに、しきい値を用いてマスク画像を生成する手法は、解像度変換 (拡大) した画像に適用すると境界線に振動が生じるなどの問題がある [7]。

2.2 SIC

SIC (Segmented Image Coding) は静止画像を領域分割し、内部情報と境界線をそれぞれ符号化する概念である。領域分割手法としては、2 乗誤差和を最小化する RSST (Recursive Shortest Spanning Tree) 手法 [10] や、自然なエッジを保持する Edgmentation 手法 [11]、境界画素数の削減に着目した改良 Edgmentation 手法 [2] などが挙げられる。

SIC において境界画素数は低ビットレートにおける符号化効率を決定するため、非常に重要である。Christopoulos らの実験によると高圧縮時には総符号量のうち 50% が境界線の符号化に費やされている [2]。そのため、ここに挙げられた領域分割手法は、均等色領域を抽出するのは当然であるが、極小領域を削減するために、値が変動するテクスチャ領域を効率的に統合することに注意が払われている。

さらに、SIC の概念に基づく領域分割ではアンチエイリアシングがほとんど考慮されていない。そのため、アンチエイリアシングを考慮した領域分割には適用できないという問題がある。

2.2.1 RSST [10]

RSST は入力画像の他に制約がないという特徴がある。領域が 4 分木のノードから構成されるという分割統合アルゴリズムなど、他のアルゴリズムは境界領域にアーチファクトが生じてしまう。さらに、RSST は領域の数を容易に制御できるため、分割画像の詳細部分を制御しやすい。

領域の評価関数として、入力画像と領域画像の 2 乗誤差の合計である SSE (Sum of Squared Error) を用いる。2 つの領域が統合されたときの SSE が、統合前の SSE の合計よりどれくらい増加するかをコストとする。このコストが最小となる領域を統合する。また、SSE を採用することによって、変換符号化と性能が比較可能になる。

本手法は、境界画素数が考慮されていないため、実際には入力画像に直接適用しても符号化効率が高くない。

2.3 Edgmentation [11]

Edgmentation は分割統合アルゴリズムの改良である。分割段階において画像の 4 分木を用いるのではなく、エッジ情報に基づいて分割を行う。そのためより自然な境界が得られる。統合は領域内の輝度の平均値や分散などを元に行う。

分割段階では、反転エッジ画像に対する最急パス成長アルゴリズム (Steepest Path Growing Algorithm) を適用する。まず反転エッジ画像から極大画素を根画素として定義する。それぞれの根画素について、入力画像に対応する画素値を与える。そして、根画素を除く任意の画素について、その画素がいずれかの根画素に到達するまで、反転エッジ画像における最急パスに沿って結合させる。同じ根画素に結合された全ての画素について、値画素と同じ画素値を与える。このようにして根画素に結合された全ての画素が小領域を構成する。

本手法は、アンチエイリアシングが適切な領域に分割される特徴を持つが、根画素の検出に性能が左右される。特に均等色領域が含まれていると根画素の検出精度が低下し、期待されるような小領域が得られない。

2.4 改良 Edgmentation [2]

Edgmentation は非常に多くの極小領域を生成する。極小領域はオブジェクトとして意味のある領域ではない。まず、平均濃度の絶対差がしきい値以下の領域群を統合する。さらに、勾配値を考慮した $CF(i, j)$ によって定義されるコスト関数を用いて統合する。 $CF(i, j)$ は領域 j を領域 i に統合する時、

$$CF(i, j) = \frac{G(i, j)}{L(i, j)} \frac{S(j)}{L(i, j)} D(i, j) \quad (3)$$

によって定義される。ここで、 $G(i, j)$ は領域 i, j の境界線に沿って勾配値を合計した値、 $L(i, j)$ は領域 i, j の共通境界線の長さ、 $S(j)$ は領域 j の画素数、 $D(i, j)$ は領域 i, j の平均濃度の差である。このコスト関数は、境界線が長く、領域が小さく、差分が小さいほど評価値が小さくなる。

提案されているコスト関数は主観品質の観点から、単独の評価値を用いたり、SSE に基づく評価値よりも、良

い結果を与える。ただし、Edgmentationによる小領域の取得が不完全な場合、後段の処理で回復できない。したがって、人工画像への適用は困難である。

2.5 Total Variation による画像分解

近年、非線形領域における数値計算手法が発展し、画像処理に適用される例が増加している。Chambolleらは、Total Variation Minimizationを用いて画像から有界変動成分を抽出する手法を提案している [12]。有界変動成分は不連続点を包含するため、画像のようにエッジを含む信号を近似することなく扱うことができ、その特徴から骨格画像と呼ばれる。なお、残差画像には主に振動成分（テクスチャ）が含まれている。

簡単のために、画像は $N \times N$ の 2 次元行列とする。 X をユークリッド空間 $\mathcal{R}^{N \times N}$ と定義する。離散化されたノルムを定義するために、離散勾配演算子を導入する。もし $u \in X$ ならば、勾配 ∇u は

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2) \quad (4)$$

となり、さらに

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & (i < N), \\ 0 & (i = N), \end{cases} \quad (5)$$

$$(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & (i < N), \\ 0 & (i = N) \end{cases} \quad (6)$$

となる。ただし、 $i, j = 1, \dots, N$ である。

u の Total Variation は

$$J(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}| \quad (7)$$

と定義される。ただし、 $|y| := \sqrt{y_1^2 + y_2^2}$ とする。

さらに、有界変動成分の抽出問題を最小化問題

$$\min_{u \in X} \frac{\|u - g\|^2}{2\lambda} + J(u) \quad (8)$$

に帰着させる (ROF モデル)。ここで、 $g \in X$ と $0 < \lambda$ は与えられた値である。 $\|\cdot\|$ は X におけるユークリッド空間のノルムである。

Total Variation の定義より、画像 u はエッジを保持する一方で振動成分が分離されている。しかし、異なる均等色領域が隣接している領域において、エッジは保存されるが強調されることはない。すなわち、領域内の均等色化が原理的に保証されていない。さらに、アンチエイリアシングも骨格画像に含まれる。

3. 提案手法

3.1 画像符号化の仕組み

人工画像をベクター表現に変換するために、SIC と同

様に領域分割を行う。人工画像ではテクスチャ領域よりもアンチエイリアシングを特に考慮する必要がある。

整数画素での領域分割では、それぞれの領域内に隣接する領域の影響が残るため、小数画素精度の領域分割を行う。すなわち、解像度変換により画像を拡大してから、整数画素精度の領域分割を適用する。得られた領域の境界線はベクター表現に変換し、領域内のテクスチャは既存の変換符号化を適用する。

3.2 逆数ノルムと目的関数

領域内における輝度の均等性を保証（均等色化）するために、隣接画素との差分が小さいほど大きい評価値となるノルムを導入する。ただし、差分がゼロの場合は、評価値（ノルム）はゼロとする。そこで、本稿では逆数ノルム

$$I(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}|^{-1} \quad (9)$$

を導入する。ただし、

$$|y|^{-1} := \begin{cases} 1/\sqrt{y_1^2 + y_2^2} & 0 \neq y_1, y_2, \\ 0 & \text{Other} \end{cases} \quad (10)$$

とする。

次に、このノルムを用いて目的関数を定義し、これを最小化する画像分解の問題に帰着させる。すなわち、

$$\min_{u \in X} \frac{\|u - g\|^2}{2\lambda} + I(u) \quad (11)$$

を解く。ここで、 $g \in X$ と $0 < \lambda$ は与えられた値である。 λ は 2 つの項を重み付けして足すためのラグランジュ係数である。

このままでは、アンチエイリアシングの除去が考慮されていない。そこで、アンシャープマスク項を導入する。具体的には目的関数を

$$\text{usm}(g) = \begin{cases} g + (g - g_s) \times k & |g - g_s| > \text{th}, \\ g & \text{Other} \end{cases} \quad (12)$$

$$\min_{u \in X} \min \left(\frac{\|u - g\|^2}{2\lambda}, \frac{\|u - \text{usm}(g)\|^2}{2\lambda} \right) + I(u) \quad (13)$$

に変更する。ここで、 $g_s \in X$ は g を空間的に平滑化したベクトル、 k はアンシャープマスクの強さを決める係数、 th はアンシャープマスクの適用を決めるしきい値である。

得られたベクトル u は、小さな振動成分が除去されて平滑化された領域が含まれており、領域間の境界が強調されている一方で、テクスチャはそのまま含まれるような画像である。

3.3 最小化アルゴリズム

提案した逆数ノルムを含む最小化問題は、数値計算的に解く方法が今のところ無い。そこで、以下に示すように、近傍の画素やアンシャープマスク後の値で置き換えた場合のコスト値を計算し、これが最小になる値を順次選択する（グリーディー法）。

まず、画像内の画素をラスタ順に処理する。現在の画素値、上にある画素値、左にある画素値、アンシャープマスク後の画素値に対して目的関数を計算する。なお、平滑化には上下左右の画素値の平均を利用する。目的関数が最小となった値を、新しい画素値として採用する。次に、逆ラスタ順に上記の手順を繰り返す。

なお、画素を評価するたびに新しい画素値に置き換える再帰的な手法と、画像全体の評価が終わった時点で全ての画素を置き換える非再帰的な手法がある。

4. 実験と考察

提案手法を実装し、計算機実験を行った。人工画像としてアニメーション画像を用いる。QVGA (320 × 240 画素) の画像 1 とこれをキュービックフィルタにより 3 倍に拡大した画像 2 を用いる。なお、パラメータとして、図 1 では $\lambda = 64$ 、図 2 では $\lambda = 572$ を用いた。また、 $th = \lambda/16$ 、 $k = 4/\lambda$ を用いた。また、最小化アルゴリズムのうち再帰法を用いた。

図 1、2 に画像 1 と画像 2 の繰り返し回数とコスト値、逆数ノルム、2 乗誤差の関係をそれぞれ示す。これらの図より、繰り返しにより逆数ノルムが低下する一方で、2 乗誤差が増加している。20 回程度繰り返したところでコスト値が最低になり、その後は 2 乗誤差の増加に伴いコスト値が上昇している。このことより、コスト値が増加を始めるところで繰り返しを終了する必要がある。

図 3 に画像 1 の 160 行目と図 4 に画像 2 の 480 行目の列座標と輝度値の関係を示す。それぞれ、提案手法による均等色化画像、Total Variation (ROF モデル) による骨格画像、入力画像を示す。これらの図より、振幅の小さい振動領域では ROF モデルも提案手法も均等色化が実現されている。図 5 に一部を拡大した様子を示す。境界領域において ROF モデルではエッジが平滑化されているのに対して、提案手法では先鋭化されていることが確認できる。このように、提案手法では逆数ノルムとアンシャープマスクを用いた目的関数により、アンチエイリアシングを考慮した均等色化が実現されている。

なお、図 6 に標準画像 Lenna の 128 行の列座標と輝度値の関係を示す。この図より、ROF モデルは柔軟な平滑化が実現されているのに対して、提案手法はテクスチャなどの比較的大きな振動を平滑化することができない。

提案した最小化アルゴリズムは、画素値を選択しているだけで平均処理などが含まれない。従って、テクスチャのように値が振動しており、かつ平均値に相当する画素が含まれていない場合、原理的に平均値に収束することはあり得ない。これはノルムではなくアルゴリズムに起因する問題である。

5. まとめ

本稿では SIC の概念に基づき、ベクトル表現による画像符号化のための人工画像を考慮した領域分割手法を提案した。まず、均等色化を実現するための逆数ノルムを提案した。さらに、アンチエイリアシングを除去するためにアンシャープマスク項を導入し、これらを含む目的関数を定義した。次に、領域分割を画像の均等色化と見なし、目的関数を最小化するアルゴリズムを提案した。計算機実験により、均等色化やアンチエイリアシングの除去が出来ていることが確認できた。今後、この出力画像を実際に均等色領域に分割し、ベクタ表現を用いた画像符号化を実現する。

謝 辞

本研究は特別研究員奨励費 (19・2363) の助成を受けたものである。

文 献

- [1] F. Marques, A. Gasull, T.R. Reed, M. Kunt, "Coding-Oriented Segmentation based on Gibbs-Markov Random Fields and Human Visual System Knowledge," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-91), Apr. 1991.
- [2] C.A. Christopoulos, W. Philips, A.N. Skodras, and J. Cornelis, "Segmented image coding: Techniques and experimental results," Signal Processing: Image Communication 11, pp.63-80, 1997.
- [3] R.L. de Queiroz, "Compressing Compound Documents," in The Document and Image Compression Handbook, edited by M. Barni, Marcel-Dekker, 2005.
- [4] K. Kawamura, H. Watanabe, and H. Tominaga, "Vector representation of binary images containing halftone dots," 2004 IEEE ICME Proceedings., Jun. 2004.
- [5] 河村, 山本, 渡辺, "文字や網点を含む低解像度多値画像のベクター表現に関する検討," FIT2006, J-032, Sep. 2006.
- [6] 河村, 山本, 石井, 渡辺, "ベクター表現による文字や網点を含む画像符号化システム," 情処研報 2006-AVM-54, no.6, Sep. 2006.
- [7] 河村, 石井, 渡辺, "ベクター表現に適した小数画素精度の改良境界線検出に関する検討," 情処研報 2007-AVM-56, no.11, Mar. 2006.
- [8] 高木, "電子書籍ビジネス調査報告書 2004," 株式会社インプレス ネットビジネスカンパニー インターネット生活研究所, 2004.
- [9] R. de Queiroz, Z. Fan, and T. Tran, "Optimizing

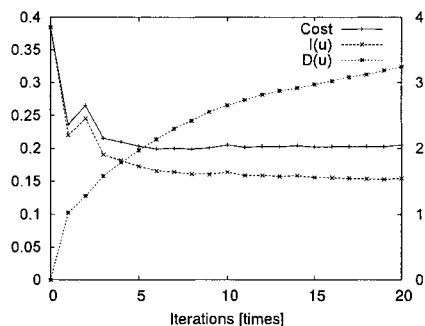


図 1 目的関数とノルムの変化 (画像 1)

Fig. 1 Transition of object function and norm (Image 1).

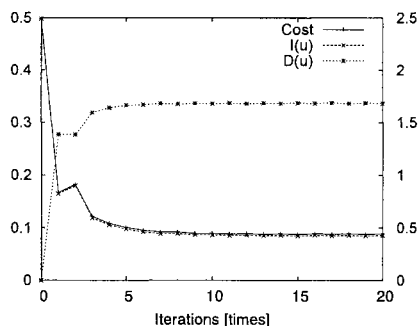


図 2 目的関数とノルムの変化 (画像 2)

Fig. 2 Transition of object function and norm.

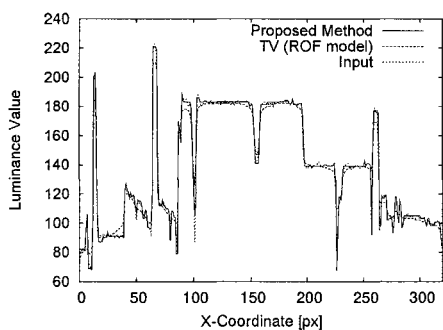


図 3 画像 1, 160 行の列座標と輝度の関係

Fig. 3 A relation of row 160 in Image 1.

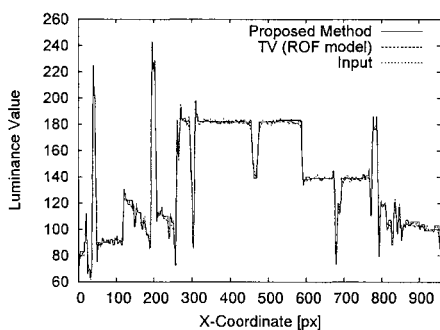


図 4 画像 2, 480 行の列座標と輝度の関係

Fig. 4 A relation of row 480 in Image 2.

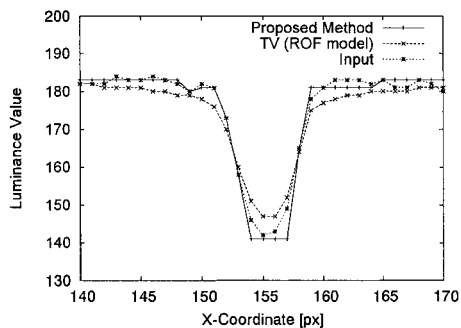


図 5 画像 1, 160 行の列座標と輝度の関係の拡大図

Fig. 5 A part of the relation of row 160 in Image 1.

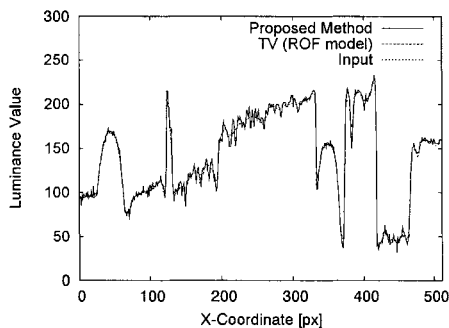


図 6 Lenna の 128 行の列座標と輝度の関係

Fig. 6 A relation of row 128 in Lenna.

block-thresholding segmentation for multi-layer compression of compound images," IEEE Trans. on Image Processing, Vol. 9, pp. 1461-1471, Oct. 2000.

- [10] M.J. Biggar, O.J. Morris, A.G. Constantinides, "Segmented-image coding: performance comparison with the discrete cosine transform," Radar and Signal Processing, IEE Proceeding F, Vol. 135, Issue 2, pp.121-132, Apr. 1988.
- [11] J. Cornelis, J.D. Becker, M. Bister, C. Vanhove, G. Demonceau, and A. Cornelis, "Techniques for Car-

diac Image Segmentation," IEEE International Conference of Engineering in Medicine and Biology Society, 1992.

- [12] A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," Journal of Mathematical Imaging and Vision 20: 89-97, 2004.