

フィールド／フレーム適応DCTを用いた 時－空間スケーラブル符号化

Spatio-Temporal Scalable Video Coding using Adaptive Infield/Inframe DCT

浅田 昌俊† 沢田 克敏‡
Masatoshi ASADA Katsutoshi SAWADA

ABSTRACT This paper describes a spatio-temporal resolution scalable coding scheme. Resolution scalability means a coding property where lower partial resolution pictures can be obtained by decoding only subsets of the total coded bit stream, while the full resolution picture is reconstructed by decoding the total bit stream. This scheme employs frame subsampling associated with adaptive interpolation for temporal scalability and adaptive infield/inframe DCT for spatial scalability. The proposed scheme provides four different spatio-temporal resolutions of a video sequence -- two temporal resolutions, each consisting of two spatial resolutions. This can be applied to interlaced video sequences effectively. Computer simulation results have demonstrated that this scheme has better coding performance compared to conventional non adaptive methods.

1. はじめに

本稿では、画像の情報を圧縮するための高効率符号化に加え、スケーラビリティを実現するためのスケーラブル符号化について検討する。スケーラビリティとは、画像データの伝送、蓄積、受信、表示などにおいて、装置の性能レベルに応じてデータの一部あるいは全部を用いて画像のサイズや解像度、動きのなめらかさなどを変更できる機能である。本研究では、時間と空間を組み合わせた時－空間スケーラブル符号化を行い、再生される時間と空間の解像度の異なる4種類の画像について検討する。その中で行われる DCT を用いた空間スケーラブル符号化では、フィールド／フレーム適応 DCT を用いる。これは、インタレース画像に対して効率よく符号化でき、空間スケーラブル符号化によって再生される空間的な低解像度画像の品質を改善することができる。また、時間スケーラブル符号化で時間的

な低解像度画像を再生するために用いるフィールド／フレーム適応内挿もインタレース画像に対して有効である。

2. 時－空間スケーラブル符号化

2.1 時－空間スケーラブル符号化の全体構成

本稿で提案する時－空間スケーラブル符号化¹⁾²⁾³⁾は、時間スケーラビリティと空間スケーラビリティを同時に実現する。そのため、時－空間スケーラブル符号化は、時間と空間のスケーラブル符号化を組み合わせた構成になっている。図1に時－空間スケーラブル符号化の構成を示す。まず、入力される画像に対して時間スケーラブル符号化を行うために奇数フレームと偶数フレームに分ける。そして、それぞれのフレームに対して DCT を用いた空間スケーラブル符号化を行い、DCT 係数を低周波成分と高周波成分に分ける。復号側では、DCT 係数の低周波成分のみを用いて空間的な低解像度画像を再生し、低周波成分と高周波成分をすべて用いて空間

† 愛知工業大学 大学院電気電子工学専攻 (豊田市)
‡ 愛知工業大学 情報通信工学科 (豊田市)

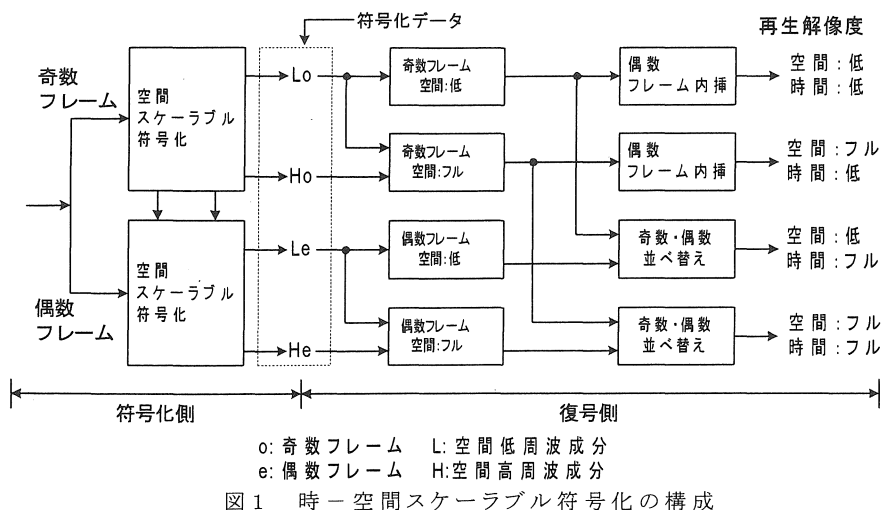


表1 符号化データと再生される画像の関係

解像度		対応するデータ
空間	時間	
低	低	Lo
フル	低	Lo, Ho
低	フル	Lo, Le
フル	フル	Lo, Ho, Le, He

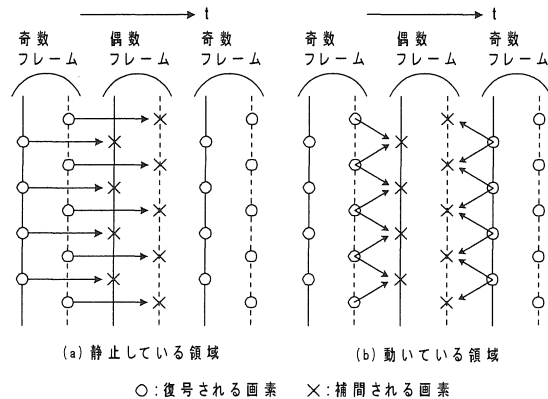


図2 フィールド/フレーム適応内挿の説明

的なフル解像度画像を再生する。また、奇数フレームのみを用いて時間的な低解像度画像を再生し、奇数フレームと偶数フレームを用いて時間的なフル解像度画像を再生する。このように、時-空間スケラブル符号化では、空間的な解像度の異なる2つの画像と時間的な解像度の異なる2つの画像を組み合わせることにより、4種類の異なる解像度の画像を再生することができる。表1に再生される4種類の画像とそれに対応する符号化データを示す。

2.2 時間スケラブル符号化

時間スケラブル符号化では、入力される画像を奇数フレームと偶数フレームに分けて処理を行う。復号側で、時間的なフル解像度画像の再生には、奇数フレームと偶数フレームの両方の符号化データを用いる。一方、時間的な低解像度画像の再生

には、奇数フレームのデータのみを用いる。まず、奇数フレームのデータを用いて奇数フレームの画像を再生し、次に、この奇数フレームの再生画像を用いて偶数フレームの画像を内挿再生する。

ここで偶数フレームを内挿するためにフィールド/フレーム適応内挿¹⁾²⁾³⁾を用いる。一般にフレーム内挿を用いる場合、静止している部分の空間解像度を保持できるが、動いている部分ではインタレース画像のため時間的な動きの順番が逆戻りする。また、フィールド内挿を用いる場合、動いている部分の時間的な順番は問題ないが、静止している部分の空間解像度は低下する。フィールド/フレーム適応内挿では、フィールド内挿とフレーム内挿の長所を生かすためにその2つの内挿方法を適応的に切り換えるようにする。すなわち図2に示すように、静

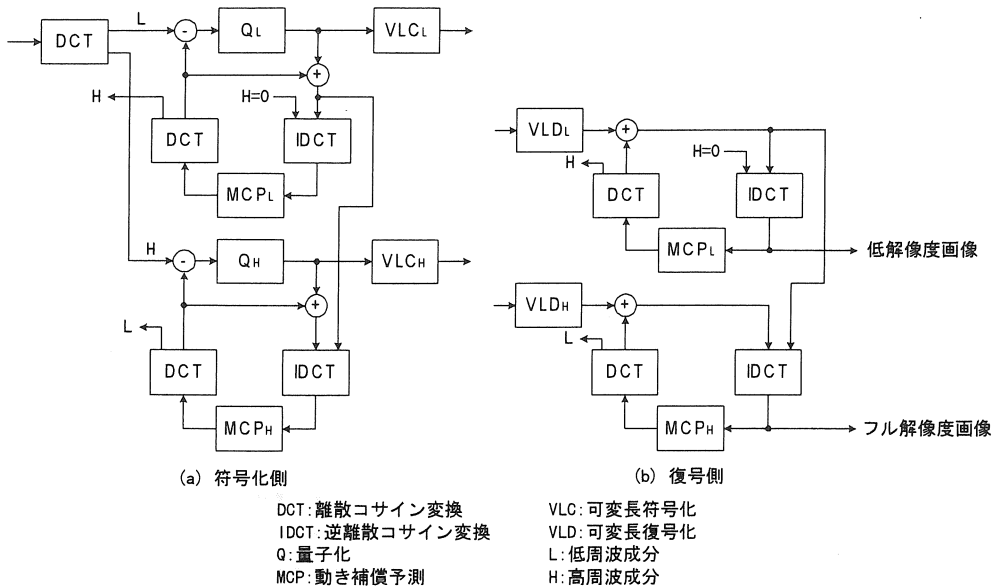


図3 MC 予測 DCT 空間スケーラブル符号化の構成

止している領域ではフレーム内挿を、また、動いている領域ではフィールド内挿を用いる。フレーム内挿では、図2(a)に示すように内挿される偶数フレームの1つ前の奇数フレームの第1フィールドと第2フィールドの画素がそれぞれ補間される偶数フレームの第1フィールドと第2フィールドの画素となる。一方、フィールド内挿では、図2(b)に示すように内挿される偶数フレームの第1フィールドは1つ前の奇数フレームの第2フィールドから空間的な画素のずれを補正するように補間され、また内挿される第2フィールドは1つ後の奇数フレームの第1フィールドから補間される。

2.3 空間スケーラブル符号化

2.3.1 MC-DCT空間スケーラブル符号化

図3に MC 予測 DCT 空間スケーラブル符号化⁴⁾⁵⁾の構成を示す。符号化側では、入力画像を8画素×8ラインに分割したブロックに対して DCT を行い、その64個の DCT 係数を低周波 (L) 成分(16個)と高周波 (H) 成分(48個)に分ける。その2つの成分は、それぞれ MC 予測符号化される。L 成分ループでは、16個の L 成分の係数と H 成分の係数を0にしたものに対して IDCT を行い画像領域に戻し

て MC 予測を行う。そして、予測画像に対して DCT を行い H 成分の係数は除去して次のフレームの L 成分との差分をとり、量子化して符号化する。一方、H 成分ループでは、48個の H 成分の係数と L 成分ループで用いられる L 成分の係数に対して IDCT を行い画像領域に戻して MC 予測を行う。そして、予測画像に対して DCT を行い L 成分の係数は除去して次のフレームの H 成分との差分をとり、量子化して符号化する。

復号側では、L 成分と H 成分のデータをそれぞれ復号し、符号化側と同じ MC 予測ループで得られた予測画像の DCT 係数との和をとる。そして、L 成分のデータのみを用いると空間低解像度画像が得られ、L 成分と H 成分の両方のデータを用いると空間フル解像度画像が得られる。

2.3.2 フィールド／フレーム適応DCT

インタレース画像を効率よく符号化するためにフィールド／フレーム適応 DCT¹⁾²⁾³⁾⁵⁾⁶⁾を用いる。フィールド／フレーム適応 DCT では、DCT を行うブロックをフィールド単位、またはフレーム単位のいずれの構成にするかを適応的に切り替えることにより、画像全体をフィールド単位やフレーム単位の構成の DCT で行うより効率よく符号化することができる。

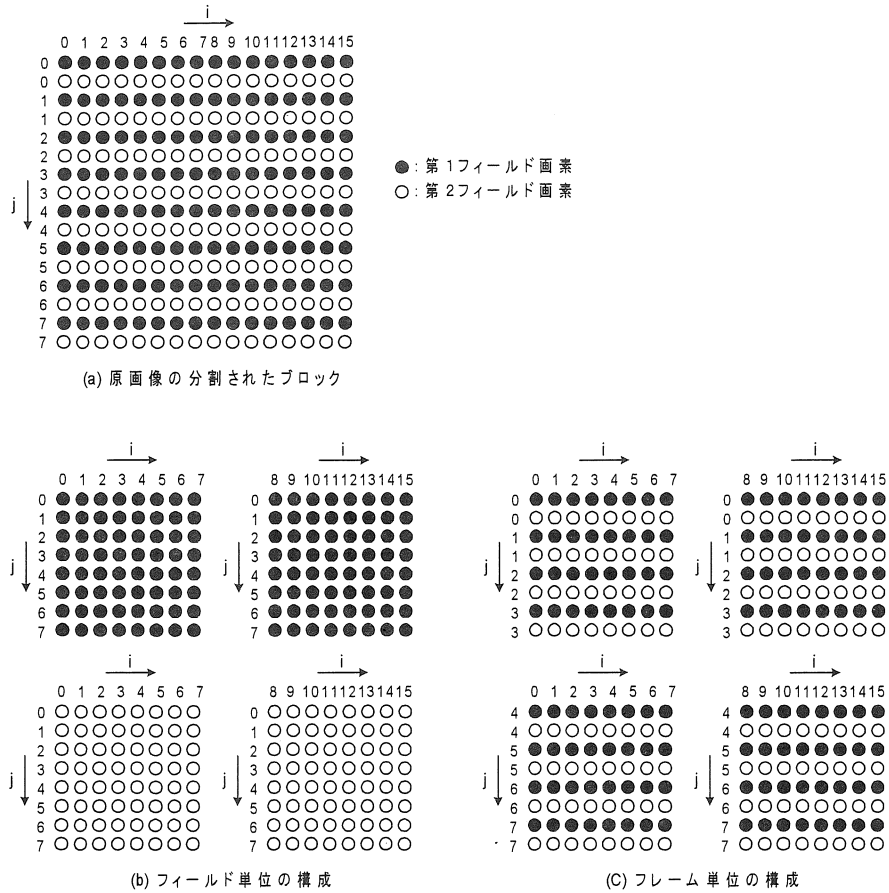


図4 フィールド/フレーム適応 DCT のブロック構成

DCT を行うブロックの構成は、図4(a)のように原画像を互いに重なり合わない16画素×16ラインのブロックに分割し、図4(b)のように4つのフィールド単位のブロックか、または図4(c)のように4つのフレーム単位のブロックに分割して8×8のDCTを行う。

DCT を行うブロックの構成を切り替える方法は、図4(a)のブロックの垂直方向の隣接する画素の相関が低いときはフィールド単位のブロック構成をし、垂直方向の隣接する画素の相関が高いときはフレーム単位のブロック構成をする。その垂直方向の相関の強さを求めるには、式(1)によって垂直方向のフィールド内差分絶対値和 $D1$ 、また式(2)によって垂直方向のフィールド間差分絶対値和 $D2$ を計算する。 $D1 \leq D2$ のときは、フィールド内での相関の方が強いのでフィールド単位でブロックを構成する。また、 $D1 > D2$ のときは、フィールド間での相関の方が強いのでフレーム単位の構成をする。

このフィールド/フレーム適応 DCT を MC 予測 DCT 空間スケーラブル符号化に取り入れることにより、再生される空間的な低解像度画像の品質を向上させることができる。

$$D1 = \sum_{i=0}^{15} \sum_{j=0}^6 (|x_{ij} - x_{i,j+1}| + |y_{ij} - y_{i,j+1}|) \quad (1)$$

$$D2 = \sum_{i=0}^{15} \sum_{j=0}^6 (|x_{ij} - y_{ij}| + |y_{ij} - x_{i,j+1}|) \quad (2)$$

x_{ij} : 第1フィールド画素値
 y_{ij} : 第2フィールド画素値

3. シミュレーション実験

3.1 実験内容

本稿で提案した時-空間スケーラブル符号化の計算機シミュレーション実験を行った。図1の構成で

時-空間スケーラブル符号化を行い、時間と空間の解像度の組み合わせによって再生される4種類の画像について検討した。また、フィールド/フレーム適応 DCT を用いた場合の効果を確認するため、フィールド DCT (画像全体をすべてフィールド単位の DCT) とフレーム DCT (画像全体をすべてフレーム単位の DCT) を用いた場合と比較検討した。実験条件を以下に示す。

・使用画像

Cheerleader, Flower garden, Table tennis

・ DCT

ブロックサイズ: 水平8画素 × 垂直8ライン

・ MC

ブロックサイズ: 水平16画素 × 垂直8ライン

探索範囲:

水平 ± 31 × 垂直 ± 15 (奇数フレーム)

水平 ± 15 × 垂直 ± 7 (偶数フレーム)

1画素精度の全探索

3・2 実験結果および考察

(1) 時間フル-空間フル解像度画像

図5は時-空間スケーラブル符号化によって得られる時間フル-空間フル解像度画像に対する符号化特性である。フィールド/フレーム適応 DCT (Adaptive DCT)、フィールド DCT (Infield DCT)、フレーム DCT (Inframe DCT) を用いた場合の3つを比較している。この図で、横軸は画像を再生するために必要な符号化エントロピーで、1画素あたりに必要な情報量(単位: [bit/pe])を表している。縦軸は原画像に対する再生画像品質を SNR (単位: [dB]) で表している。

この図から、3つの DCT の差は同じエントロピーで SNR を比較したとき 1[dB]以内であることがわかる。実際に再生した画像を観察した結果、3つの DCT の違いによる差はほとんど確認できなかった。したがって、時間フル-空間フル解像度画像の場合には、符号化特性と実際に再生画像を観察した結果から DCT の違いによる大きな差はほとんどないと言える。

(2) 時間フル-空間低解像度画像

図6は時-空間スケーラブル符号化によって得られる時間フル-空間低解像度画像に対する符号化特性である。この図から同じエントロピーで SNR を

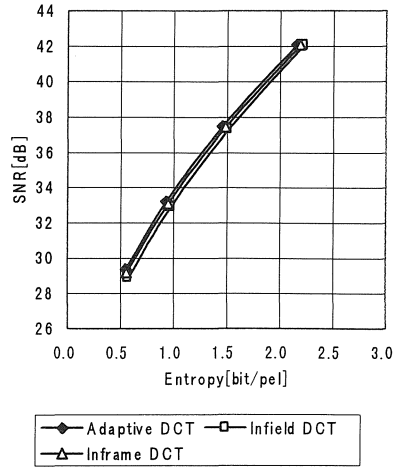


図5 時間フル-空間フル解像度画像の符号化特性 (Cheerleader)

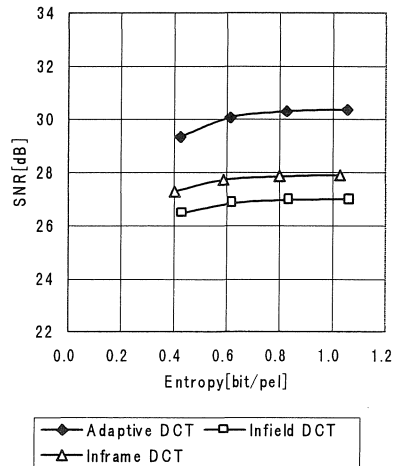
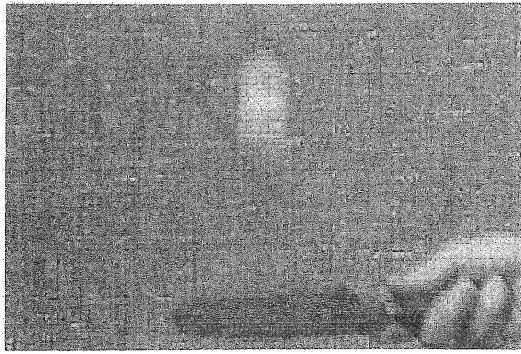
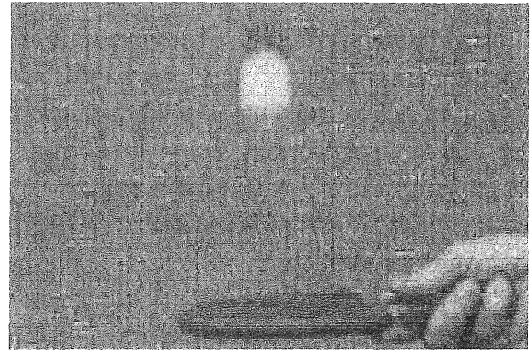


図6 時間フル-空間低解像度画像の符号化特性 (Cheerleader)

比較したとき、フィールド/フレーム適応 DCT を用いた場合、フィールド DCT やフレーム DCT を用いたときより最高で 3.5[dB]程 SNR が高いことがわかる。実際に再生した画像を、図7、図8に示す。図7のように動いている領域でフレーム DCT を用いると画像がぼやけてしまうが、フィールド/フレーム適応 DCT を用いた場合には良好な画像が得られる。また、図8のように静止している領域(家の屋根の部分)でフィールド DCT を用いると画質劣化が大きいが、フィールド/フレーム適応 DCT を用いるとフィ

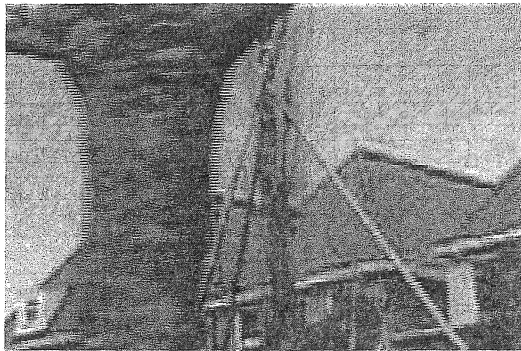


フレーム DCT

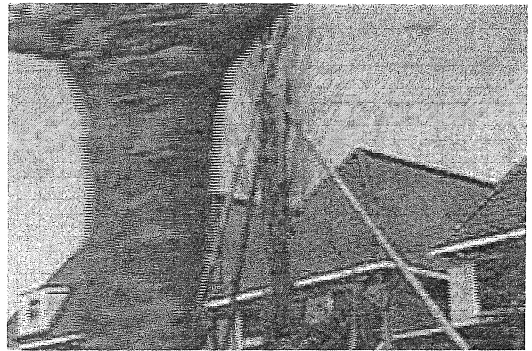


フィールド/フレーム適応 DCT

図7 空間低解像度画像の動いている領域の例



フィールド DCT



フィールド/フレーム適応 DCT

図8 空間低解像度画像の静止している領域の例

ールド DCT を用いた場合より画像品質が良くなる。つまり、フィールド/フレーム適応 DCT を用いた場合、フィールド DCT やフレーム DCT を用いた場合に比べ画像品質が向上することがわかる。このように、符号化特性と実際に再生画像を観測した結果からフィールド/フレーム適応 DCT の効果が確認された。

(3) 時間低-空間フル解像度画像

時間低-空間フル解像度画像は、符号化データから復号される奇数フレームを用いて偶数フレームを内挿している。そのため、時間低-空間フル解像度画像を再生するために必要な符号化エントロピーは、時間フル-空間フル解像度画像の場合の約半分になる。再生画像を観察した結果、偶数フレームの内挿にフィールド/フレーム適応内挿を用いたため、フレーム内挿のような動き順番の逆転の問題はなく、また、フィールド内挿のような静止部分の空間解像度の低下も生じなかった。

(4) 時間低-空間低解像度画像

時間低-空間低解像度画像も、符号化データから復号される奇数フレームを用いて偶数フレームを内挿している。そのため、時間低-空間低解像度画像を再生するために必要な符号化エントロピーは、時間フル-空間低解像度画像の場合の約半分になる。再生画像を観察した結果、この場合もフィールド/フレーム適応内挿の効果が確認された。

4. むすび

本稿では、フィールド/フレーム適応 DCT とフィールド/フレーム適応内挿を用いた時-空間スケラブル符号化について検討した。計算機シミュレーション実験の結果、時-空間スケラブル符号化を行うことにより、復号側で時間と空間の解像度の異なる4種類の画像を再生することができた。また、インタレース画像に対してフィールド/フレーム適応 DCT を用いることにより、フィールド DCT やフレーム DCT に比べ、空間的な低解像度画像の品質

を改善することができた。さらに、フィールド/フレーム適応内挿の効果も確認することができた。

文献

- 1) 浅田, 沢田: "フィールド/フレーム適応 DCT とフレーム内挿を用いた時-空間解像度階層符号化", 1996 年電子情報通信学会総合大会 D-260, 1996.3.
- 2) 浅田, 沢田: "フィールド/フレーム適応 DCT を用いた時-空間解像度階層符号化", 1996 年テレビジョン学会年次大会 23-6, 1996.7.
- 3) M.Asada and K.Sawada: "A Scalable Video Coding Scheme Based on Adaptive Infield/Inframe DCT and Adaptive Frame Interpolation", Proc. IWISP'96, pp.257-260, Nov. 1996.
- 4) M.Nakamura and K.Sawada: "Scalable Coding Schemes based on DCT and MC Prediction", Proc. ICIP'95, vol.2, pp.575-578, Oct. 1995.
- 5) 浅田, 沢田: "インタレース画像に対する動き補償予測 DCT 階層符号化", 1995 年電気関係学会東海支部連合大会 553, 1995.9.
- 6) 八島, 上倉, 沢田: "適応ブロック構成 DCT を用いた HDTV-標準 TV コンパチブル符号化", 1989 年電子情報通信学会秋季大会 D-54, 1989.

(受理 平成9年3月21日)