

顕著度を考慮した多視点画像特徴量を用いた三次元形状類似検索 3D model retrieval using salient visual features from multiple viewpoints

長田 邦男¹, 坂野 智久², 大渕 竜太郎³
Kunio Osada¹, Tomohisa Banno², Ryutarou Ohbuchi³

^{1,2,3} 山梨大学 〒400-8511 山梨県甲府市武田 4-3-11
^{1,2,3} University of Yamanashi, 4-3-11 Takeda, Kofu-shi, Yamanashi-ken, Japan

¹g06mk005@yamanashi.ac.jp, ²t01k073f@ccn.yamanashi.ac.jp, ³ohbuchi@yamanashi.ac.jp

1. はじめに

3次元モデルを形の類似性で比較する際には、まずその形をコンパクトに表現する特徴量を求め、比較したいモデル同士で特徴量の比較を行う。特徴量は、形状全体から抽出する大域特徴量と、形状の複数の部分から抽出する局所特徴量に分けることができる。局所特徴量の利点は、例えば人モデルの姿勢変化（手足や腰の屈曲など）のように、大域特徴の変化は大きい局所特徴の変化は小さい変形には比較結果が影響されにくいことである。反面、その数が多くなりすぎると相違度の計算コストが上昇する。

Shilane ら[7]は、3次元形状モデル1つを数千個の領域に分割し、それぞれの領域から求めた形状特徴を局所特徴量とした。彼らは学習フェーズにおいて、検索に有効な局所特徴量を学習し、それに基づいて、数千個の局所特徴量の中から検索に有効なものを3~10個程度選別し、選別した局所特徴量の集合を3次元モデルの特徴量とする手法を提案した。彼らの実験から、3次元モデルの局所特徴量には、限られた数の有意な局所特徴量を用いることにより、計算コストと検索性能を両立できることが分かった。

我々は、3次元モデルの見かけ（複数の2次元画像）を元に、それら2次元画像の局所特徴を組み合わせて3次元形状の類似比較を行う手法を提案する。画像の中で重要な点を選択して特徴を求めるSIFT法[1]を用いることにより、局所特徴数を抑えつつ、検索性能を高めることができた。評価実験の結果、SHREC 2006[2]に参加した17手法中で、SHREC2006における検索性能の評価尺度の1つFirstTier(%)で1位と同等という高い性能を得た。

2. 提案手法

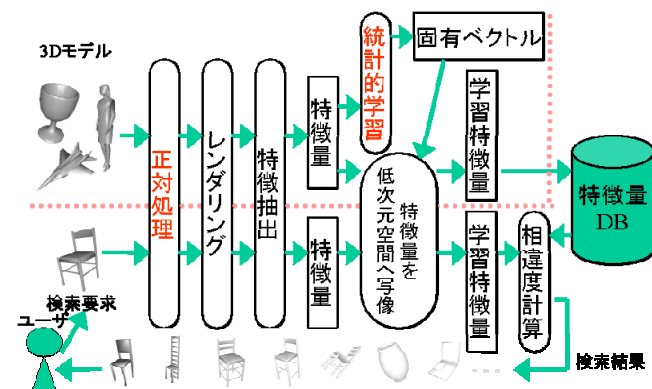


図1: 本検索手法の処理工程。

提案手法では、一般的な三次元形状類似検索の処理工程に正対処理（モデルの向き合わせ）と特徴量の統計的学習を加えた。提案手法の処理工程は図1のようになる。

まず、同一視点での類似比較をするために、3次元形状モデルの大きさ、位置、向きを正規化する正対処理を行い、得られた直行3軸の正負の向き合計6方向から並行投影により視点からの深さ画像（z値画像）を求める。6個のz値画像それぞれについてLoweらのSIFT法[1]により顕著点の検出、および顕著点での局所画像特徴量（SIFT特徴量）の抽出を行う。ここで、求めた特徴量に対し統計的学習手法に基づく次元削減を適用し、特徴量の識別性を高める。

2.1. 多視点レンダリング

今回、我々は3次元モデルの視点からの距離画像（z値画像）からSIFT特徴量を得る方法を用いた。投影方法には正射影を用い、視点はx, y, zの直行3軸上の正負の方向からの計6視点とした（図2参照）。このようにして得た6枚のz値画像からSIFT特徴量を抽出し、それらの集合を3次元モデルの特徴量とする。ただし、2.5節で述べるように相違度計算は同一視点のSIFT特徴量群の対を用いる。

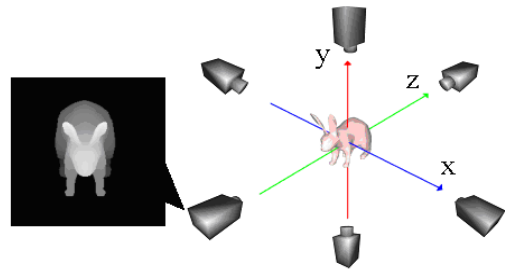


図2: 直行3軸上正負の方向、計6視点から撮影。

2.2. 3次元モデルの正対処理

我々が評価実験に用いるSHREC2006で検索対象とするデータベースでは3次元モデルの向きや大きさ、位置が正規化されていない。従って、非常に類似した形状を持つ2つのモデルをレンダリングした画像、ひいてはこれらの画像から計算したそれぞれのSIFT特徴量群が大きく異なる可能性がある。本手法では、3次元モデルの大きさ、向き、位置の正規化（正対処理）を行うことで、この問題を緩和した。具体的には、Point SVDとNormal SVDと呼ばれる2種の正対処理[3]を組み合わせる。図3はこれら2種の正対手法を適用した例で、特にNormal SVDが優位である場合を示したものである。

Point SVD: モデルの表面にばら撒いた点群の座標値の分布

を主成分分析(PCA)して求めた3つの主軸を3次元モデルの固有3軸とすることで正対処理を行う。PCAによる正対処理では主軸の正負は定まらない。そこで、モデルの固有3軸に沿った Bounding Box の中心を原点とし、そのどちら側にモデルの重心があるかで固有3軸の正負を定める。モデルの重心はモデルの面上に生成した点群の座標値の平均とする。また、3次元モデルの重心を原点とすることで位置の正規化を行い、ばら撒いた点群の内、3次元モデルの重心から最遠点となる点までの距離を1とすることで大きさの正規化を行った。

Normal SVD : モデルの表面にばら撒いた向き付き点群の法線ベクトルの分布を主成分分析して求めた3つの主軸を3次元モデルの固有3軸とする正対処理。モデルの面に撒いた点の法線ベクトルは、その点がかかる面の法線ベクトルとする。大きさの正規化や軸の正負の決定は Point SVD と同様に行う。

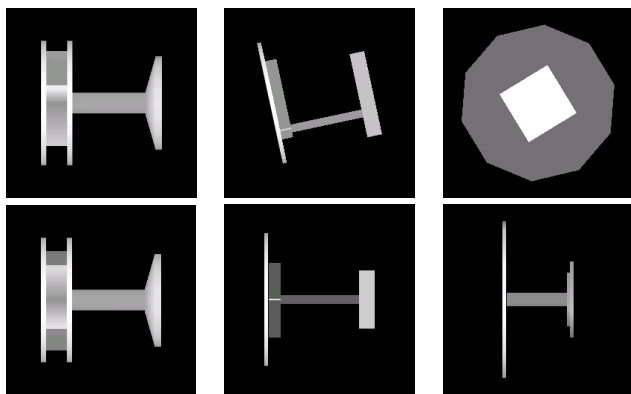


図 3 : SHREC2006 に使用されるテーブルのモデルに対して Point SVD(上段)と Normal SVD(下段)を適用した例。

2.3. 特徴量抽出

SIFT 法[1]は Lowe によって提案された画像特徴抽出手法で、画像から「人が気にするような点」(顕著点)を複数検出し、そのような点周辺(局所)の画素値に基づいた特徴量を抽出する。つまり、SIFT 特徴量は画像用の局所特徴量であり、顕著点の数は画像ごとに異なる。SIFT 特徴量は画像の回転、拡大縮小、アフィン変形、照明変化、ノイズ付加に頑強であるといった特性を持つ。

SIFT 法は人の視覚に基づいていると言われており、3次元形状の類似比較を、人の視覚に基づいて行うのに SIFT 特徴量を用いるのが適していると考え、使用した。

顕著点には以下(a)~(d)の属性が与えられる。

- (a) 座標値
- (b) スケール
- (c) 向き
- (d) SIFT 特徴量

このうち(a)は二次元画像中の顕著点の位置を、(b)は顕著点の特徴領域の大きさを、(c)は顕著点の特徴領域の回転方向を捕らえる向きを、それぞれ示す。(d)は上述したように顕著点の周辺の特徴を128次元ベクトルで示している。本手法では(d)だけでなく、(a)も相違度計算に用いることで検索性能を向上させた。なお、SIFT 特徴量の抽出には、Lowe により Web ページに公開されている実行形式のプログラムを利用した。図4は異なる人型の3次元モデル2つのz値画像から SIFT 法で求めた顕著点の例である(z軸の正の方

向からの1視点のみ)。手、脚、頭、首下、胴体など、人が見て共通した部分に顕著点が現れていることがわかる。

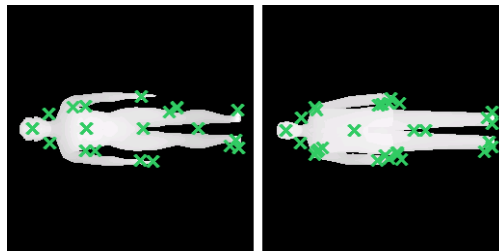


図 4 : 人型のモデル2つのz値画像から SIFT 法で求めた顕著点の例を×印で示す。(Point SVD を適用済み)

2.4. 相違度計算

SIFT特徴量を用いたモデル間の相違度計算は次のようにして行う。Point SVDを1, Normal SVDを2とする正対処理手法Nを適用したとき、モデルiの特徴量の集合を $X_{N,i}$ 、総視点数を n 、モデルiの視点vから得られた特徴量を $x_{N,i,v}$ 、モデルiの視点vからレンダリングした画像を $I_{N,i,v}$ 、画像 $I_{N,i,v}$ から SIFT特徴量を得る関数を $sift(I_{N,i,v})$ とすると、特徴量 x_i は以下のようにして求まる。

$$X_{N,i} = \{x_{N,i,0}, x_{N,i,1}, \dots, x_{N,i,n}\} \quad \dots(1)$$

$$x_{N,i,v} = \{sift(I_{N,i,v})\}$$

ここで注意して欲しいのは、 $x_{i,v}$ はSIFT法がその画像から検出した顕著点の数だけあり、その数が画像ごとに(したがって3次元モデルごとにも)異なる点である。

次に、検索要求モデルQと検索対象のモデルDの相違度を求める関数 $Distance(X_q, X_d)$ を定義する。

モデルQとDの特徴量をそれぞれ X_q, X_d 、2種の正対処理手法Nをそれぞれ適用したときに得る2つの距離のうち、小さい方の距離を選択する関数を $\min_N$ 、視点vにおけるモデルQとDの相違度を求める関数を

$distance(X_{N,q,v}, X_{N,d,v})$ とし、 $Distance(Q, D)$ を式(2)で表す。

$$Distance(X_q, X_d) = \min_N \left(\sum_{v=1}^n distance(X_{N,q,v}, X_{N,d,v}) \right) \quad \dots(2)$$

提案手法では、関数 $\min_N$ によって、2種の正対手法の組み合わせを行った。

上述したように、顕著点の数は3次元モデルごと、視点ごとに異なるため、距離計算には工夫が必要である。本手法では3次元モデルに対して正対処理を行ってあるため、似たような3次元モデル同士であれば、空間的に近い位置に似た特徴が現れる可能性が高い。そこで、顕著点の属性値、(a)座標値に基づいて、画像上の顕著点の空間近傍を取ること、距離比較をする顕著点の対を絞る処理を行った。

モデルQの視点vにおいて、顕著点の総数をsとし、任意の顕著点pのSIFT特徴量を $x_{N,q,v,p}$ 、モデルDの視点vにおける顕著点の中から、(a)座標値に基づいてpとのk近傍を取ったとき、それらの持つSIFT特徴量との最小距離を求める関数を $\min_k(x_{N,q,v,p}, X_{N,d,v})$ とし、 $distance(X_{N,q,v}, X_{N,d,v})$ を式(3)で表す。

$$distance(X_{N,q,v}, X_{N,d,v}) = \frac{1}{s} \sum_{p=1}^s \min_k(x_{N,q,v,p}, X_{N,d,v}) \quad \dots(3)$$

また, 顕著点の空間近傍の他, 距離の計算の際に以下の工夫を行った.

顕著点数による比較の足切り: 顕著点数に大きな開きがあること自体, その1対のモデル間の相違度が大きいことを示す場合が多い. そこで, 視点 v における顕著点数の差が一定値 t を超えた場合には, 式(3)を計算せずに視点 v における距離 $distance(X_{N,q,v}, X_{N,d,v})$ を無限大とするようにした. この処理を足切りと呼ぶ.

足切りが発生したモデルDの救済: 足切りを行うかどうかは, 視点 v における顕著点数の差が一定値 t を超えるかどうかで決定される. このため, 中にはモデルQとモデルDが似ているにもかかわらず, 足切りが発生してしまう場合がある. このような場合の救済措置を目的として, 視点 v において足切りが発生した場合の視点 v における距離 $distance(X_{N,q,v}, X_{N,d,v})$ を求める. 検索要求モデルQに対して足切りが発生しなかったデータベースモデルをモデルUとし, モデルQと全てのモデルUとの相違度の集合を求める関数を $Distance(X_q, U)$ とする. $Distance(X_q, U)$ の内, 最も大きい値を求める関数を $\max(Distance(X_q, U))$ としたとき, 足切りが発生した視点 v における距離 $distance(X_{N,q,v}, X_{N,d,v})$ を式(4)とする.

$$distance(X_{N,q,v}, X_{N,d,v}) = \max(Distance(X_q, U)) / n \quad \dots(4)$$

2.5. 統計的学習による SIFT 特徴量の次元削減

各 SIFT 特徴量に対して次元削減を行い, 計算量の削減, 及び, 検索性能の向上を試みた. 我々は, Ke [4]が SIFT 特徴量の次元削減に用いた PCA の他, 独立主成分分析 (ICA), と Locality Preserving Projections (LPP) [5]の合計 3 つの手法を比較した. 最適な削減後の次元数は手法ごとに異なるはずであるが, 今回は, 3 手法全て, PCA の寄与率が 99% となる 103 次元に次元削減した. 学習に際して, データベースの全モデルから得られる SIFT 特徴量の総数は正対処理手法ごとに約 50 万個だが, これらのインデックスをコンテナに入れてランダムソートし, 先頭から 8000 個をサンプルし, 正対処理手法ごとに学習した. ランダムソートには MinGW, g++, STL の random_shuffle 関数を用いた.

3. 実験と結果

検索性能評価実験は SHREC2006 ベンチマークに従った. SHREC 2006 ではポリゴンスーパーモデルから成る PSB (Princeton Shape Benchmark) [6]の test set モデル 907 個と train set モデル 907 個を合わせた計 1814 個のモデルを検索対象データベースとする. 検索要求はデータベースに含まれない 30 個のモデルである. ここでは各パラメータの探索のための検索性能の基準となる尺度として FirstTier (Highly Relevant クラス)の値 [%]を用いた. FirstTier は情報検索における R-Precision と同意である. 以降, 特に断りが無い限り, 検索性能とは FirstTier (Highly Relevant クラス)のこととし, 単に FT と表記する.

まず, 実験学習手法ごと, 距離尺度ごとの検索性能の実験を行った. その実験に従い, 検索性能が良かった距離尺度を用いて, 学習手法ごとに足切りの実験を行った. 同様にして, 式(3)における k の値の実験を行い, 正対処理手法の組み合わせの実験と同時に足切り救済措置の実験を行っ

た. 最後に実行時間に関する実験を行った.

なお, z 値画像の画像サイズについては, 今回は 256×256 画素とした.

3.1. 学習手法

各正対手法ごと, 各学習手法ごとに, 3 種の距離尺度 (マンハッタン, ユークリッド, コサイン距離) を比較した. 結果は表 1 の通りである. Original は学習なしの場合であり, これと比較すると, 両正対手法において, どの学習手法を用いても検索性能が向上することが分かった. 特に, モデルに Point SVD を適用し, LPP で学習した場合には 7% 以上という非常に有意な性能向上が見られた. なお, 学習手法によらず, Point SVD を適用したときの検索性能は Normal SVD を上回っていることから, Point SVD を適用したほうが, Normal SVD を適用したときよりも, 全体的に人が期待するような正対が行われているのだと考えられる.

以降の実験では, 学習手法ごとに最も検索性能が高かった距離尺度を用いることとする. ただし, LPP に関しては, Point SVD ではマンハッタン距離, Normal SVD ではユークリッド距離を用いたときに最高性能を示したため, 正対手法ごとに両距離尺度を用いて実験を行った.

3.2. 足切り

各正対手法ごと, 各学習手法ごとに, 最も良い検索性能となる足切りの閾値 t の探索を行った. 表 2 にその結果を記す (LPP で学習し, 距離尺度をマンハッタン距離としたものを LPP_Man, LPP で学習し, 距離尺度ユークリッド距離としたものを LPP_Euc と表記する). これにより, 全体として t は 50 付近の値を用いることでほぼ最高性能となることがわかった. ただし, Point SVD を用い, SIFT 特徴量を LPP で学習し, マンハッタン距離を用いたときでは $t=120$ と, 他に比べ t が大きな値のときに最高性能となる. また, LPP は他の学習手法に比べ足切りによる性能の伸び幅が小さく, この時点で PCA が LPP を上回る結果となった.

以後の実験では t の値を, 学習手法ごと, 正対処理手法ごとに最も性能が良かった値を使用する.

3.3. 顕著点の空間近傍

式(3)における k の値を決定する実験を行った. 正対手法ごと, 特徴量の学習手法ごとに, 最高性能を得た k の値を表 3 に記す. この中で最も良い性能を出したのは, Point SVD を用い, PCA で学習した場合で, k を 24 としたときで, FT は 42.17% あった. これは k 近傍を用いなかった場合 (表 2) よりも約 4 ポイントと, 大きな性能向上が見られた. Normal SVD を用い, PCA で学習した場合も約 3% ポイントの性能向上が見られた. これに対し, ICA, LPP で学習した場合の表 2 からの性能向上値は 0~1.8 ポイント程度で, 大きな性能向上は見られず, また k の値も PCA に比べ大きかった. そして, この時点で, ICA, LPP で特徴量を学習しても, 検索性能が学習無しとほとんど変わらないか, 学習無しが上回ってしまう結果となった.

3.4. 正対手法の組み合わせと足切りの救済

3.3 節までの実験によって決定された学習手法ごとの距離尺度, パラメータ t と k を用いて, 式(2)による正対手法の組み合わせを行った結果と, 各正対手法ごとに式(4)で定義した足切りの救済処置を行った上で, 正対手法の組み合わせを行った結果を表 4 に記す. 足切りの救済処置を行わ

なかったときには, 学習なしを除いては, 表3における Point SVD の結果から 2 ポイント程度の性能向上が見られた. そして, 足切りの救済措置を適用したとき, 特徴量を PCA で学習したときの検索性能は 44.96% となり, SHREC2006 での 1 位, Makadia らによる 44.77% を上回る結果となった. PCA, ICA, LPP で学習したときの足切りの救済措置による性能向上はわずかであった (一部は減少した) が, 学習なしの場合には 1.5 ポイント以上の性能向上がみられ, 全体としては足切りの救済措置を行ったほうが良い結果を得られるのではないかと考えられる.

3.5. 実行時間

実験には Athlon64 3500+ (シングルコア), メモリ 1GB のマシンを使用した. ここでは主に相違度計算の実行時間について議論する. 相違度計算の実行時間は, 検索要求モデル 1 個, データベースモデル 1814 個あたりに掛かった時間とする. 3.5 節の実験で最高性能を示した, 特徴量を PCA で学習した場合では, Point SVD で約 4.4 秒, Normal SVD で約 4.8 秒であった. なお, 特徴量を PCA で学習し, 足切りをせず, 顕著点の空間近傍も考慮しなかった場合は Point SVD で約 7.39 秒, Normal SVD で約 7.14 秒であった. したがって, Point SVD では約 40%, Normal SVD では約 33% 秒, 正対手法を組み合わせる場合は約 58%, 相違度計算時間を削減できたことになり, 足切り, 顕著点の空間近傍を考慮することで, 相違度計算時間を大幅に削減しつつ, 検索性能を向上できることがわかった. なお, 相違度計算以外の処理を含めた, 検索要求 1 つあたりのシステムの実行時間は約 16.4 秒であった (正対処理に約 2.6 秒, 多視点画像生成に約 1.3 秒, 特徴量抽出に約 3.2 秒, 正対手法の組み合わせに 0.05 秒)

4. まとめと今後の課題

本研究では, 3次元形状の比較に同モデルの見かけ2次元画像から求めた局所特徴量を用いる方法を提案し, その有効性を示した. LoweらのSIFT法[1]を用いて検出した顕著点についてのみ局所特徴を求めることで, 計算量を抑えることに成功した. さらに, モデルの表面に生成した点群の座標値と法線ベクトルを考慮することで2種類の正対処理を行い, SIFT特徴量の次元削減, 顕著点数差による足切り, 比較する顕著点の空間近傍に基づく限定, などの工夫により検索性能を向上させた. その結果, SHREC 2006[2]に参加した17手法と比較し, 評価尺度にFirstTier(%)を用いたとき, 1位と同等という高い性能を得た.

今後, 各種パラメータを適応的に変更させることや, より良い特徴量の学習手法の探索, そして, より良い相違度計算アルゴリズムの考案などが必要である. 計算時間についても, 検索システムの実用的な範囲に収めるために更に改善が必要である.

5. 参考文献

- [1] David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, *Int'l Journal of Computer Vision*, 60(2), November 2004.
- [2] R. C. Veltkamp, et. al, SHREC2006 3D Shape Retrieval Contest, Utrecht University Dept. Information and Computing Sciences, Tech. Report UU-CS-2006-030 (ISSN: 0924-3275)
<http://give-lab.cs.uu.nl/shrec/shrec2006/index.html>

- [3] 立間 淳司, 関 洋平, 青野 雅樹: 多重フーリエスペクトル表現に基づく 3 次元モデルの形状類似検索, 第 7 回 Web インテリジェンスとインタラクション研究会, 電子情報通信学会, pp.89-94, 東京大学科学技術研究センター, 11 月 12~13 日, 2006
- [4] Y. Ke, R. Sukthankar: PCA-SIFT: A more distinctive representation for local image descriptors, *Proc. CVPR2004*.
- [5] Xiaofei He, Partha Niyogi: Locality Preserving Projections, *Advances in Neural Information Processing Systems 16*, Vancouver, Canada, 2003.
<http://people.cs.uchicago.edu/~xiaofei/LPP.html>
- [6] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton Shape Benchmark. *Proc. SMI 2004*.
- [7] Philip Shilane and Thomas Funkhouser. Selecting Distinctive 3D Shape Descriptors for Similarity Retrieval. *IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06)*

表 1: 正対手法, 学習手法, 距離尺度ごとの検索性能

measure	Point SVD			
	Original	PCA	ICA	LPP
Manhattan	29.88	35.69	32.01	37.40
Euclidean	27.77	27.62	35.07	37.30
Cosine	22.10	21.89	32.53	33.05
measure	Normal SVD			
	Original	PCA	ICA	LPP
Manhattan	28.39	29.90	25.94	30.86
Euclidean	24.67	24.20	30.19	31.21
Cosine	19.41	19.06	27.56	28.31

表 2: 最高性能を出した t の値とそのときの検索性能

	Point SVD				
	Original	PCA	ICA	LPP_Man	LPP_Euc
t	40	50	50	120	40
FT(%)	34.29	38.61	36.78	37.69	38.03
	Normal SVD				
	Original	PCA	ICA	LPP_Man	LPP_Euc
t	60	60	60	50	40
FT(%)	30.30	33.20	33.56	32.84	33.08

表 3: 最高性能を出した k の値とそのときの検索性能

	Point SVD				
	Original	PCA	ICA	LPP_Man	LPP_Euc
k	19	24	24	28	30
FT(%)	39.13	42.17	37.28	39.34	39.83
	Normal SVD				
	Original	PCA	ICA	LPP_Man	LPP_Euc
k	24	24	100	100	70
FT(%)	33.89	36.15	33.56	32.84	33.30

表 4: 正対手法の組み合わせ結果

(足切り救済措置の有無での性能を比較)

	足切り救済措置なし				
	Original	PCA	ICA	LPP_Man	LPP_Euc
FT(%)	39.76	44.55	39.20	41.24	40.74
	足切り救済措置あり				
	Original	PCA	ICA	LPP_Man	LPP_Euc
FT(%)	41.43	44.96	38.95	40.98	41.03