# クラスタリングに基づく情報の検索と視覚化

堀田，政二

https://doi.org/10.15017/1398256

出版情報：九州芸術工科大学，2001，博士（工学），課程博士

# 付 録B

# 2次形式距離

2次形式距離 (Quadratic-Form Distance) はヒストグラム間の類似性を反映した距離尺度であり，ヒストグラムを直接比較するミンコフスキー距離 (Minkowski-Form Distance) よりも視覚特性に沿った距離関係を与える [14, 44]．そのため QBIC[1]ではカラー画像検索に応用されている．本付録ではカラーヒストグラム間の2次形式距離の定義を概説し，2次形式距離を利用して画像を低次元空間に配置する方法 [45] を示す．

## B.1　カラーヒストグラム間の2次形式距離

はじめに，画像がカラーヒストグラムで与えられる場合の画像間の2次形式距離の定義を示す．画像 $i$ の $n$ 色のカラーヒストグラムを $\boldsymbol{h}_i = [h_{i1}, ..., h_{in}]^T$ とする．$\sum_{r=1}^{n} h_{ir} = 1$ となるように規格化しておく．画像 $i$ と $j$ の2次形式距離は

$$\begin{aligned} QFD^2(\boldsymbol{h}_i, \boldsymbol{h}_j) &= (\boldsymbol{h}_i - \boldsymbol{h}_j)^T S(\boldsymbol{h}_i - \boldsymbol{h}_j) \\ &= \sum_{p=1}^{n}\sum_{q=1}^{n}(h_{ip} - h_{jp})s_{pq}(h_{iq} - h_{jq}) \end{aligned} \tag{B.1}$$

によって定義される [14, 44]．ここで $S$ は色 $p$ と $q$ の類似度 $s_{pq}$ を要素とする $n \times n$ の実対称行列である．$S = [s_{pq}]$ を色類似度行列と呼ぶ．2次形式距離は $S$ の設定によって負の値をとる可能性があり，その場合には距離の公理を満たさない．そこで文献 [14, 44] では，行列 $S = [s_{pq}]$ を $s_{pq} = 1 - d_{pq}/d_{\max}$ とする方法が提案されている．ここで $d_{pq}$ は色 $p$ と $q$ の Luv 空間でのユークリッド距離であり，$d_{\max} = \max_{pq} d_{pq}$ である．本論文では，より知覚特性に沿った距離を得られるように，色 $p$ の Lab 値を

[1] http://wwwqbic.almaden.ibm.com/

$L_p, a_p, b_p$ として $s_{pq}$ を

$$s_{pq} = e^{-\alpha[(L_p - L_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2]} \tag{B.2}$$

によって設定している．ここで $\alpha$ はファジー度を調節するパラメータであり，この値が小さいほど色 $p$ と $q$ の類似度は大きくなる．

次に，2 次形式距離がミンコフスキー距離よりも知覚特性に沿った距離関係を与えることを簡単な例で示す．例として赤，オレンジ，青からなる画像があり，それぞれのカラーヒストグラムを $\boldsymbol{h}_r = [1, 0, 0]^T$, $\boldsymbol{h}_o = [0, 1, 0]^T$, $\boldsymbol{h}_b = [0, 0, 1]^T$ とする．色類似度行列 $S$ を

$$S = \begin{bmatrix} 1.0 & 0.9 & 0.0 \\ 0.9 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \tag{B.3}$$

とすれば，2 次形式距離の大小関係は $QFD(\boldsymbol{h}_r, \boldsymbol{h}_o) \leq QFD(\boldsymbol{h}_r, \boldsymbol{h}_b) = QFD(\boldsymbol{h}_o, \boldsymbol{h}_b)$ となる．一方，画像間の距離をミンコフスキー距離

$$D_{L_p}(\boldsymbol{h}_i, \boldsymbol{h}_j) = \left( \sum_{r=1}^{n} |h_{ir} - h_{jr}|^p \right)^{\frac{1}{p}} \quad (p \geq 1) \tag{B.4}$$

によって計算すると $D_{L_p}(\boldsymbol{h}_r, \boldsymbol{h}_o) = D_{L_p}(\boldsymbol{h}_r, \boldsymbol{h}_b) = D_{L_p}(\boldsymbol{h}_o, \boldsymbol{h}_b)$ となり，互いの距離はすべて等しくなる．この例から 2 次形式距離がミンコフスキー距離よりも人間の知覚に近い距離関係を与えることが容易に理解できる．

## B.2　2 次形式距離を利用した低次元空間への画像配置

本節では，色類似度行列 $S$ の固有値分解に基づいて，画像を低次元ユークリッド空間に配置する方法を示す．この方法を応用したナビゲーションによる画像検索の例も示す [45].

### B.2.1　特徴ベクトルの次元削減

色類似度行列 $S = [s_{pq}]$ の固有値分解を $S = U\Lambda U^T$ とする．$U = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]$ は固有ベクトル $\boldsymbol{u}_i$ を並べた直交行列，$\Lambda$ は固有値を並べた対角行列である．すると式 (B.1) は $QFD^2(\boldsymbol{h}_i, \boldsymbol{h}_j) = (\boldsymbol{h}_i - \boldsymbol{h}_j)U\Lambda U^T(\boldsymbol{h}_i - \boldsymbol{h}_j)$ となり，ヒストグラム $\boldsymbol{h}_i$ を

$$\boldsymbol{x}_i = \Lambda^{1/2} U^T \boldsymbol{h}_i \tag{B.5}$$

により $\boldsymbol{x}_i$ に変換すると，式 (B.1) は $\boldsymbol{x}_i$ のユークリッド距離の 2 乗

$$
\begin{aligned}
D^2(\boldsymbol{x}_i, \boldsymbol{x}_j) &= (\boldsymbol{x}_i - \boldsymbol{x}_j)^T(\boldsymbol{x}_i - \boldsymbol{x}_j) \\
&= \sum_{r=1}^{n}(x_{ir} - x_{jr})^2
\end{aligned}
\tag{B.6}
$$

となる．この標準形の上位 $l \leq n$ だけの $D_l^2(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{r=1}^{l}(x_{ir} - x_{jr})^2$ で $D^2(\boldsymbol{x}_i, \boldsymbol{x}_j)$ を近似したのが低次元距離であり $D_l^2(\boldsymbol{x}_i, \boldsymbol{x}_j) \leq D^2(\boldsymbol{x}_i, \boldsymbol{x}_j)$ が成り立つ．

この次元削減を実際に行うときには，行列 $S$ は個々の画像には関係なく共通なので，検索に先立ってあらかじめ固有値分解しておき，各画像の $n$ 次元ヒストグラム $\boldsymbol{h}_i$ を $l$ 次元ベクトル $\boldsymbol{x}_{il} = \Lambda_l^{1/2} U_l^T \boldsymbol{h}_i$ に変換しておく．ここで $\Lambda_l$ と $U_l$ は $\Lambda$ と $U$ において値が大きい $l$ 個の固有値とそれに対応する固有ベクトルだけを残したものであり，$\Lambda_l^{1/2}$ は固有値の平方根を要素とする対角行列である．この $\boldsymbol{x}_{il}$ が画像の $l$ 次元空間での座標である．この $l$ 次元空間はユークリッド空間である．すなわち，距離 $D_l(\boldsymbol{x}_i, \boldsymbol{x}_j)$ は $\boldsymbol{x}_{il}$ と $\boldsymbol{x}_{jl}$ のユークリッド距離で求められる．したがって，$l$ 次元空間の表示は通常の直交座標でよい．

### B.2.2 実験例

大きさ $150 \times 150$ の風景写真 500 枚を使って実験を行った．すべての画像を 512 色に減色した．また，色類似度行列 $S$ を式 (B.2) で作成した．$\alpha = 0.001$ とした．$l$ を 2 として画像を配置したものを図 B.1 に示す．配置にかかった計算時間は 0.01 秒であった．後述する MDS やスプリングモデルによる配置は中心付近に画像が多く固まる性質があり，画像を弁別し難いが，本節の配置法ではそれよりも均等に分布しており，個々の画像を見やすい．また，従来の多くの配置法ではデータが追加された場合，新たにデータ全体の配置を計算し直す必要があるが，本節の配置法では追加データの座標を式 (B.5) で計算して追加表示するだけでよい．したがって，例えばインタネットで画像を集めてきて表示する場合，従来法では全部の画像データが集まるまで配置表示できないが，本節の配置法では画像を取得するごとに順番に画像を配置表示していくことができる．

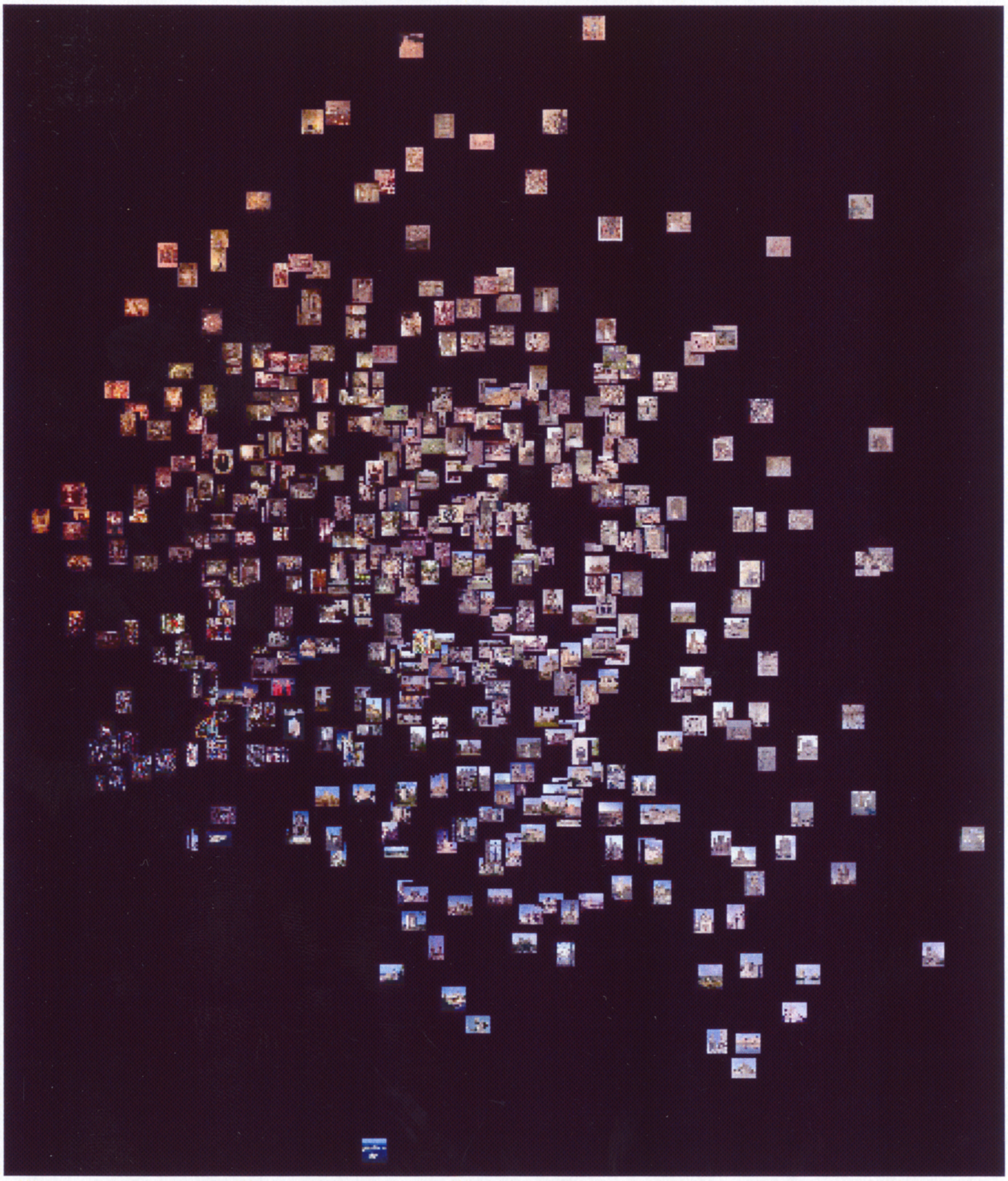

図 B.1: 2 次形式距離による画像の配置

### B.2.3 低次元空間でのナビゲーションによる画像検索への応用

筆者らは前節の画像配置法を利用して，画像の $k$NN 探索を支援する視覚化システムを提案した [45]．一般に，画像検索システムでは $k$NN 検索が行われ，クエリに近い $k$ 個の画像が選ばれて表示され，ユーザは表示された $k$ 個の画像の中から次のクエリを選ぶというのを繰り返して所望の画像を探していく．通常，$k$ 個の画像は単に画面上に並べて表示され，それらの互いの類似関係はわかり難い．また，この探索プロセスではクエリの局所近傍が次々に表示されるだけなので，データベース全体の中の大域的な検索の道筋もわかり難い．

そこで，2 つのウインドウを用いて，片方に高次元で検索された $k$ 個の画像を逐次に表示し，他方にデータ全体中の検索履歴の道筋を表示することによって，$k$NN 探索プロセスの進行を支援する方法を提案した．2 次形式距離を利用した画像の配置において，各画像の座標は特徴値だけで決まるので，計算が高速であり，データ集合が変わってもデータの位置関係は不変なので，データベース全体での整合性が保たれる．したがって，データベース全体の表示空間の中にそれまでたどってきた探索の道筋を表示することにより，現在の位置と向きを常時確認することができる．図 B.2 に提案手法による画像探索システムの表示例を示す．左のウインドウでは，クエリ (四角で囲まれている) と，512 次元でクエリに近い画像上位 9 個が 3 次元空間に配置されている．また，右のウインドウにはデータベースの全画像が 3 次元空間に配置表示され，現時点までの検索パスが折線で示されている．両ウインドウとも VRML で表示されており，各々独立にズームや回転などが行える．ユーザが左の画面で 10 個のなかのどれか 1 つの画像をクリックすると，クリックされた画像が新たにクエリとなり，左ウインドウは 512 次元でそのクエリに距離が近い上位 9 個の画像とクエリの表示に更新される．このとき右のウインドウでは選ばれた新しいクエリに 1 ステップ前のクエリから線が引かれる．図 B.2 はこのステップを 15 回繰り返したときの画面である．

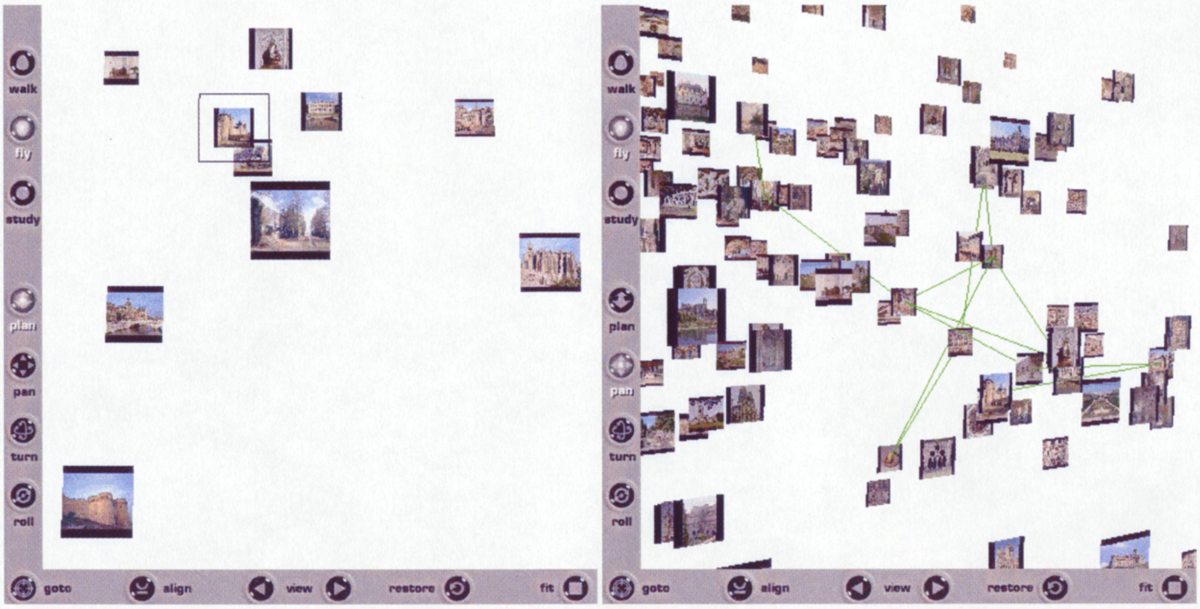


図 B.2: $k$NN 検索のナビゲーションシステム