

乳房 X 線写真上の腫瘍陰影検出システムにおける類似画像検索法を用いた偽陽性削除法

栗山 浩一[†]中川 俊明^{††}原 武史[†] (正員)藤田 広志[†] (正員)遠藤登喜子^{†††}岩瀬 拓士^{††††}

False-Positive Elimination for Mass Detection System on Mammograms Using Image Retrieval Approach

Koichi KURIYAMA[†], Toshiaki NAKAGAWA^{††}, *Nonmembers*,
Takeshi HARA[†], Hiroshi FUJITA[†], *Members*,
Tokiko ENDO^{†††}, and Takuji IWASE^{††††}, *Nonmembers*[†] 岐阜大学大学院医学研究科再生医科学専攻知能イメージ情報部門, 岐阜市
Department of Intelligent Image Information, Division of
Regeneration and Advanced Medical Science, Graduate
School of Medicine, Gifu University, 1-1 Yanagido, Gifu-
shi, 501-1193 Japan^{††} 岐阜大学バーチャルシステムラボラトリー, 岐阜市
Virtual System Laboratory, Gifu University, 1-1 Yanagido,
Gifu-shi, 501-1193 Japan^{†††} 国立名古屋病院放射線科, 名古屋市
Department of Radiology, National Hospital of Nagoya,
4-1-1 Sannomaru, Naka-ku, Nagoya-shi, 460-0001 Japan^{††††} 癌研究会附属病院, 東京都
Cancer Institute Hospital, 1-37-1 Kamiikebukuro,
Toshima-ku, Tokyo, 170-8455 Japan

あらまし 乳房 X 線写真上の腫瘍陰影の検出システムにおける偽陽性の削除を目的とした類似画像検索法を利用した手法を提案する。本手法では、画像の類似性を局所パターンマッチングから得られる特徴量によって判定する。腫瘍陰影と偽陽性陰影各 53 枚、及び偽陽性陰影 1526 枚を用いた二つの実験から、本手法の有用性を示した。

キーワード 乳房 X 線写真, 腫瘍陰影, 類似画像検索, 局所パターンマッチング, 偽陽性削除

1. ま え が き

CAD (Computer-Aided Diagnosis) システムは、コンピュータを用いて医用画像の解析及び定量化を自動的に行い、第 2 の意見として医師にその結果を提示することにより、読影診断の支援を行うことを目的としている。我々は乳房 X 線写真 CAD システムの開発を行っており、本システムを構成する要素の一つである腫瘍陰影の検出システムにおいては、偽陽性の削除が大きな課題となっている。最近、病変の良悪性の鑑別に対するコンピュータの分析結果を医師に提示する際、従来のように数値で示すのではなく、診断の対象としている画像と類似する画像をデータベースから検索し、それらの診療情報を医師に提供する方式が提案されている。医師がこれらの確定診断がなされてい

る類似画像を参照することにより、診断の精度を向上させることが期待できる。我々は、その目的のもと、腫瘍陰影の検出システムの開発と、その検出性能の向上を目指し、これまでに検出法と偽陽性の削除方法をいくつか報告してきた [1]~[4]。同時に、画像検索法についても報告を行ってきた [5]。画像検索の手法を腫瘍陰影の偽陽性削除法へ応用する研究は非常に少なく、我々の知る限り Li らによる手法 [6] のみである。本研究は、これまでに我々が提案した局所的なパターンマッチングによる画像検索の手法 [5] を、乳房 X 線写真上の腫瘍陰影検出システムにおける偽陽性削除処理に応用するものである。つまり、腫瘍陰影検出システムで指摘された領域と類似する画像を、あらかじめ用意した腫瘍陰影と偽陽性陰影の画像データベースから検索し、その結果から候補領域が腫瘍陰影か偽陽性陰影かを判定する。中川らの手法 [5] には、使用する特徴量の選択が容易に行えること、原画像の画素値から直接計算でき簡便であること、あらかじめ画像の領域分割や特徴量と類似画像との相関関係の学習の必要がないことなどの利点がある。

医用画像においては、領域分割が困難であり、医師でなければキーワードを与えることができない場合が多い。また、現在の腫瘍陰影の検出システムの性能では、偽陽性の個数が多く、実用化のためには、偽陽性の削除による性能向上が必要とされている。

本研究では、局所的なパターンマッチングによって得られる画像の特徴空間を主成分分析を用いて次元の削減を行い、ユークリッド距離に基づいた類似度から類似画像を検索したのち、偽陽性の削除判定を行う手法を提案する。そして、腫瘍陰影と偽陽性陰影各 53 枚、及び偽陽性陰影 1526 枚を用いた二つの実験から本手法の有用性を検討した。ここで、腫瘍陰影とは検出システムが正しく腫瘍として検出した真陽性の陰影のことであり、偽陽性陰影とは検出システムが正常領域を誤って腫瘍陰影と判断した陰影のことである。

2. 方 法

2.1 前 処 理

局所パターンを用いる場合、画像のサンプリング間隔と階調数の設定は、結果に大きな影響を与える。そのため、検索の対象によって、サンプリング間隔、画素値及び階調数を変換する前処理が必要である。対象である腫瘍陰影は、ある程度の大きさ (約 5 mm 以上) をもつことからダウンサンプリングを行っても問題ないと考えた。そこで、8×8 画素の画素値の平均値を

1 画素値とすることによりサンプリング間隔 $400\mu\text{m}$ の画像を利用することとした。そして、候補ごとに異なる濃度のばらつきを補正するため、領域内のコントラストを保ったまま、領域内の最大画素値が 4095 となるように画素値の変換を行った。その後、特徴抽出に用いる局所パターンを抑制 (568 個) するために、階調数を原画像の 4096 階調から 64 階調に変換する処理を行った。

2.2 特徴抽出

偽陽性削除処理の対象となる画像 (以下、質問画像) 及び検索の対象となるデータベース内の画像 (以下、検索対象画像) に対して特徴抽出を行う。特徴抽出は、図 1 に示すような局所パターンが画像全体に存在する個数を調べ、これを特徴量とする。図 1 において、各画素の濃淡は画素値に対応する。そして、局所パターンは画素数によって分類され、それをレベルと呼ぶ。例えば、レベル 2 の局所パターンは 2 画素から構成され、各画素の画素値及び結合方向の組合せだけ異なるものができる。これらのパターンの種類には無数の可能性があるが、画素値は対象となる画像の階調の範囲で用意する。

各画像に対して、 n 種類の局所パターンを用いて特徴抽出を行うと、 n 種類の要素をもつ特徴量が得られる。ここで用いた局所パターンは、最大 2 画素で構成される。レベル 2 の局所パターンについては、画素間の濃度差が 1 である 504 種類、レベル 1 の局所パターンについては対象画像の階調数を 64 階調としたので 64 種類となり、合計 568 種類の局所パターンを用いることになる。

2.3 主成分分析を用いた検索特徴空間の作成

候補領域の面積の違いが特徴量の大きさに影響を与えるため、特徴量をレベル 1、レベル 2 でそれぞれ平均が 0、分散が 1 となるように正規化した。その後、 n 次元の特徴量から検索に有効な特徴を反映した m ($< n$) 次元空間 (以下、検索特徴空間) を作成するため主成分分析を行い、主成分を軸とした検索特徴空間

間を作成する。画像の検索特徴空間における特徴量の各要素は、主成分得点とした。主成分分析を用いた次元数の縮小については、累積寄与率を変化させたときの腫瘍陰影と偽陽性陰影のそれぞれの残存率を調べ、最適な条件を求めた結果から、主成分の累積寄与率が 75% 以上となるように次元を選択した。累積寄与率が 75% のとき、次元が 45 となったため、特徴抽出で得られた 568 次元の特徴量を 45 次元に縮小して検索特徴空間を作成した。

2.4 類似度の算出

作成した検索特徴空間において、質問画像と検索対象画像の特徴量間のユークリッド距離 D を式 (1) から算出し、これを類似度と定義する。

$$D = \sqrt{\sum_{k=1}^m (Z_{ak} - Z_{bk})^2} \quad (1)$$

ここで、 Z_{ak} , Z_{bk} はそれぞれ画像 a , b の第 k 主成分得点である。質問画像とすべての検索対象画像間で類似度を算出し、各質問画像について検索結果画像を決定した。判定に用いる検索結果画像の枚数を変化させたときの腫瘍陰影と偽陽性陰影のそれぞれの残存率を調べ、最適な条件を求めた結果から、検索結果画像枚数は、質問画像の特徴量との距離に近い 8 枚で、そのうち偽陽性陰影の画像枚数が 7 枚以上のとき、質問画像を偽陽性と判定した。

3. 結果と考察

スクリーン/フィルム系で撮影された乳房 X 線画像をサンプリング間隔 $50\mu\text{m}$ 、階調数 4096 階調でデジタル化した画像データベースにおいて、候補領域を切り出した 106 枚 (腫瘍陰影 53 枚、偽陽性陰影 53 枚) を利用した。候補領域は、我々の腫瘍陰影検出アルゴリズム [1], [2] で検出された陰影を含む最小限の方形領域とした。図 2 にその例を示す。腫瘍検出アルゴリズムで検出された陰影の辺縁が灰色の線で示されており、

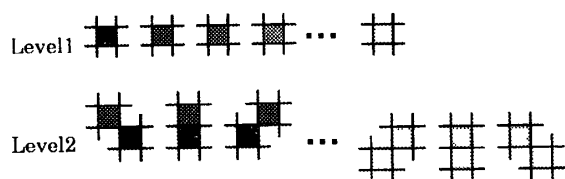


図 1 特徴抽出に用いる局所パターン
Fig. 1 Local patterns for feature extraction.

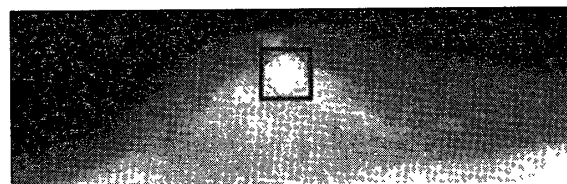


図 2 腫瘍陰影を含む最小方形による候補領域の抽出
Fig. 2 Extraction of a candidate region by rectangular window which has a minimum area including a mass shadow.

Query images	Retrieved images	Number of masses	Number of false-positive
		8	0
		0	8

図3 質問画像とその検索結果画像の例（黒枠：腫瘍陰影，枠なし：偽陽性陰影）
Fig. 3 Examples of query images and retrieved images. (black flame: mass, no flame: false-positive)

その領域を含む最小限の方形（黒線）で候補領域を抽出している。ただし、腫瘍陰影検出アルゴリズムで抽出した候補の大きさや形状が、医師の描いたスケッチと異なる陰影がいくつか存在した。そこで、すべての腫瘍陰影に対して、動的輪郭モデルを用いた辺縁抽出法[7]を適用し、より正確な腫瘍陰影を含む候補領域において特徴量を抽出した。

106枚の画像から質問画像を1枚抽出し、他の105枚の画像との間で類似度を算出し、検索結果画像を決定した。図3に検索結果の例を示す。左に質問画像を示し、その右に類似画像8枚を示した。黒枠の画像が腫瘍陰影を含む画像であり、枠なしの画像は偽陽性陰影を含む画像である。また、検索結果画像8枚中に含まれる偽陽性陰影の枚数を0から8枚まで変化させ、削除判定を行ったときの腫瘍陰影と偽陽性陰影のそれぞれの残存率を図4に示す。腫瘍陰影の質問画像53枚に対して53枚（100%）を正しく腫瘍陰影であると判定し、偽陽性陰影の質問画像53枚に対して45枚（85%）を正しく偽陽性であると判定することができた（図4の①）。図5に偽陽性の質問画像を誤って腫瘍陰影であると判定した8枚の画像とその検索結果画像、及び検索結果画像における腫瘍陰影と偽陽性陰影の各枚数を示す。図3と同様に、黒枠の画像が腫瘍陰影を含む画像であり、枠なしの画像は偽陽性陰影を含む画像である。これら8枚の偽陽性陰影には、比較的円形度が高く厚い乳腺が塊を形成している陰影が多く見られた。偽陽性陰影は、乳腺や血管などの領域であるが、一見して腫瘍陰影との類似度が低い索状の陰影や、類似度の高い比較的円形度が高く厚い乳腺が塊を形成しているように見える陰影などが存在した。そのため、実験では索状の陰影など腫瘍陰影との類似度が低い偽陽性陰影は比較的容易に削除できることが分

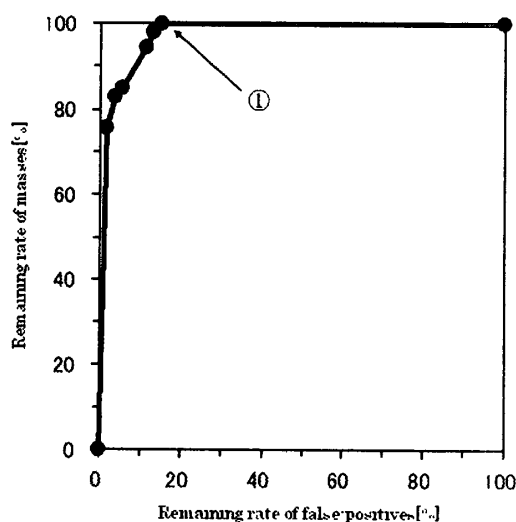


図4 腫瘍陰影の残存率と偽陽性陰影の残存率
Fig. 4 Relationship between remaining rate of masses and false-positives.

かった。しかしながら、円形度が高く腫瘍陰影との類似度が高い偽陽性陰影は、検索結果画像に多くの腫瘍陰影を含んでしまい、削除できないことが分かった。これは、検索対象画像のデータベースが106枚と比較的小規模で、種類に富んでいなかったためと考えられる。

次に、大規模な画像群への適用例として、これまでの実験で用いた106枚の画像群を検索対象画像データベースとし、新たな偽陽性陰影の質問画像1526枚について実験を行った。その結果、1526枚のうち1124枚（74%）を正しく偽陽性陰影であると判定し削除することができた。この結果から、腫瘍陰影の検出システムの性能向上が示唆できた。

4. むすび

局所的なパターンマッチングによる画像検索の手法を乳房X線写真上の腫瘍陰影検出システムにおける偽

Query images	Retrieved images	Number of masses	Number of false positive
		8	0
		2	6
		3	5
		6	2
		4	4
		4	4
		4	4
		7	1

図 5 腫瘍陰影であると誤判定した偽陽性の質問画像とその検索結果画像（黒枠：腫瘍陰影，枠なし：偽陽性陰影）

Fig. 5 Query images of false-positive detected as a mass, and the retrieved images. (black frame: mass, no frame: false-positive)

陽性削除処理に応用した手法を提案した。腫瘍陰影と偽陽性陰影各 53 枚，計 106 枚を用いた実験では，腫瘍陰影を一つも削除することなく，約 85% の偽陽性陰影を削除することができた。また，偽陽性陰影 1526 枚のデータベースを用いた結果では，その約 74% の偽陽性陰影を正しく削除することができた。これらの結果から本手法がマンモグラム CAD システムにおける腫瘍陰影の偽陽性削除法として有用である可能性を示した。

文 献

- [1] 松原友子，藤田広志，遠藤登喜子，堀田勝平，池田 充，木戸長一郎，石垣武男，“乳房 X 線写真における腫瘍陰影検出のためのしきい値法に基づく高速処理アルゴリズムの開発”，Med. Imaging Technol., vol.15, no.1, pp.1-13, 1997.
- [2] 松原友子，笠井 聡，関 和泰，藤田広志，原 武史，遠藤登喜子，“マンモグラムのためのコンピュータ診断支援システムの開発—腫瘍陰影の自動検出における低濃度領域抽出法の改善”，日乳癌検診学会誌，vol.7, no.1, pp.87-101, 1998.
- [3] 笠井 聡，松原友子，原 武史，藤田広志，畑中裕司，遠藤登喜子，“マンモグラム上の乳腺領域周辺に存在する腫瘍陰影検出に特化したアルゴリズムの開発”，医用電子と生体工学，vol.38, no.2, pp.111-117, 2000.
- [4] Y. Hatanaka, T. Hara, H. Fujita, S. Kasai, T. Endo, and T. Iwase, “Development of an automated method for detecting mammographic masses with a partial loss of region,” IEEE Trans. Med. Imaging, vol.20, no.12, pp.1209-1214, 2001.
- [5] 中川俊明，原 武史，藤田広志，“局所的なパターンマッチングによる画像検索法”，信学論 (D-II), vol.J85-D-II, no.1, pp.149-152, Jan. 2002.
- [6] Q. Li, S. Katsuragawa, and K. Doi, “Computer-aided diagnostic scheme for lung nodule detection in digital chest radiographs by use of a multiple-template matching technique,” Med. Phys., vol.28, no.10, pp.2070-2076, 2001.
- [7] 中川俊明，原 武史，藤田広志，岩瀬拓士，遠藤登喜子，堀田勝平，“乳房 X 線画像における腫瘍陰影の辺縁抽出法”，信学技報，MI2002-84, 2003.

(平成 15 年 4 月 1 日受付，7 月 9 日再受付)