

医薬品パッケージの類似性の 検出と可視化

関谷卓[†] 渡辺秀文^{††} 齋藤隆文[†]

現在、医薬品の取り違えという医療ミスが問題となっている。この原因の一つに、医薬品の見た目が類似していることが挙げられる。医薬品は膨大な量があり、注意すべき医薬品を把握することは困難である。そこで本研究では、医薬品の類似性を検出、可視化できるシステムを提案する。医薬品には錠剤を用い、錠剤のパッケージ画像から色や印字のパターンなどの特徴をもとに類似性を測る。可視化には直感的に関係がわかるよう、二次元グラフを用いる。今回は特に、錠剤パッケージに印字された文字の類似性の検出手法を検討する。

Detection and Visualization of Similarity for Medicine Packages

Taku Sekiya[†], Hidefumi Watanabe^{††}
and Takafumi Saito[†]

Mistaking medicines is one of the possible medical errors. One reason of the mistake is that there are medicines similar in look. It is hard to know such similar medicines in advance because there are a large amount of medicines. We propose a system which detect and visualize the similarity of tablet medicines. We determine the similarity of colors, printed patterns and so on from package images of tablets. In visualization, we use the two dimensional graph to show the relations between medicines. In this report, we especially focus on detection methods for the similarity of printed characters.

1. はじめに

現在、医療の現場では医薬品の取り違えという医療ミスが問題になっている。医薬品の取り違えの原因として、医薬品の名称が類似していることや、処方する患者を間違えてしまうことなどがあるが、本研究では医薬品パッケージのデザインが類似していることによる取り違えについて注目する。取り違えを防ぐために、類似した医薬品は保管する棚の位置を離しておく、などの対応[1]がなされているが、医薬品には膨大な量があるために手作業で類似したものを探し出すことは困難である。また、製薬会社が新たな医薬品を作り出しても、そのデザインが既存の医薬品のデザインと類似してしまう可能性がある。

医薬品のパッケージデザインには色や文字などの様々な特徴がある。このような特徴に関して、それぞれの医薬品でどれだけ類似しているのかを測ることが出来れば、どの医薬品が類似しているのかを知る手がかりとなる。また、新たに作り出した医薬品のデザインが、既存の医薬品のもので類似しているものがあるかどうかを知る手がかりとなる。そこで本研究では、医薬品パッケージの特徴を検出、可視化し、膨大な量の医薬品を比較できるシステムの提案を行う。

2. 類似性の検出と定量化

医薬品にはカプセルや錠剤、粉末、液体など様々なものがあるが、本研究では錠剤に着目する。錠剤の特徴について、検出手法や定量化手法を述べる。

2.1 錠剤の特徴

実際に医薬品を区別する際、例えばパッケージの色や印刷された文字、医薬品そのものの違いなどを見て判断すると考えられる。そこで、本研究で扱う錠剤について、そのパッケージデザインの特徴について述べる。パッケージの画像は各製薬会社からダウンロードすることができる。パッケージの画像は図 1(a)のように裏と表が撮影されており、本研究では特にパッケージの表側の特徴について扱う。パッケージの特徴として、ここでは以下の特徴について取り扱う。

[†] 東京農工大学大学院 生物システム応用科学府
Graduate School of Bio-Applications and Systems Engineering,
Tokyo University of Agriculture and Technology, Japan

^{††} 東京農工大学 工学部
Faculty / Graduate School of Engineering, Tokyo University of Agriculture and Technology

- ・ 背景色
- ・ 印刷された文字の色
- ・ 印刷された文字の大きさ
- ・ 印刷パターンによる文字の繰り返し
- ・ 錠剤の色
- ・ 錠剤の大きさ
- ・ 錠剤の数

図1(b)にこれらの特徴を示す。以降では、これらの7つの特徴について扱う。

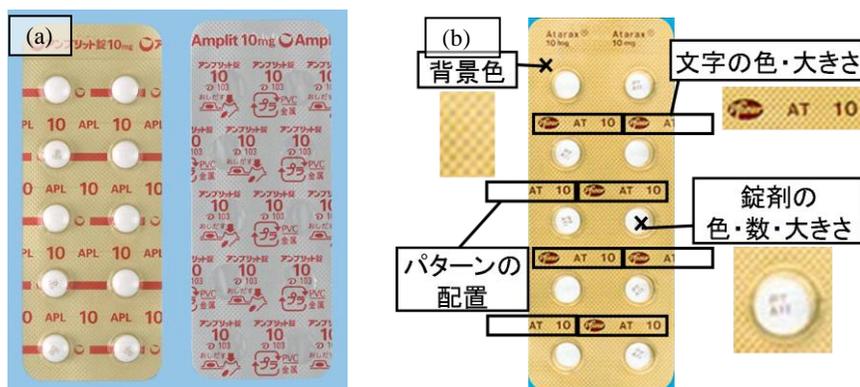


図1 錠剤パッケージの画像と特徴

Figure 1 The example of the tablet package image.

2.2 類似性の検出

類似性の検出では、錠剤の画像を図2(a)のように背景、文字、錠剤の三つの領域に分ける。このような特定領域の抽出にはJ.R. Smithらによるタグ付けを用いた色領域の抽出[2]や、色とテクスチャの抽出[3]といった手法がある。本研究では、明るさや色相といった色の違いを用いて、錠剤を三つの領域に分ける。これらの分離された画像から色についての情報や、文字、錠剤についての情報を取得することができる。

次に、分離した画像のうち文字だけの画像を用いて、文字の繰り返しを検出する。繰り返しのパターンを検出する手法としては、Liu, Y.らによる帯状模様と壁紙模様に基づく繰り返しパターン検出[4]がある。本研究では、図2(b)のように文字だけの画像から、文字やマークで構成されるパターンとなる部分を一つ任意に選択し、テンプレートマッチングを行い、同じパターンを検出する。このようにして、文字やマークがどのように繰り返しているのかを検出する。

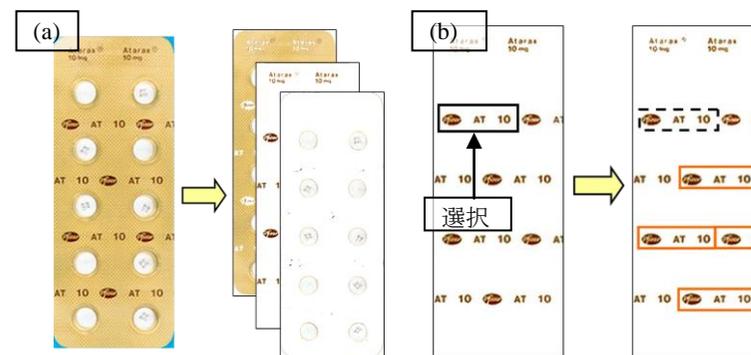


図2 繰り返しパターンの検出

Figure 2 Detection of the tablet package.

2.3 類似性の評価と定量化

検出した特徴を比較し、可視化するために定量化を行う。それぞれの特徴についてどのように定量化を行うのかについて述べる。

(1) 色

背景色、文字色、錠剤の色について定量化を行う。色の違いによって分離された画像はRGB値で表されている。これらをCIE-L*a*b*表色系に変換する。ここからL*a*b*空間上での距離である色差と、色相を求めて定量化する。

(2) 文字の大きさ

文字の大きさは、文字部分の面積の割合をを求めることで定量化する。文字だけに分離された画像を用いて、パッケージ画像における文字部分の画素数を数え、画像に対する割合を求める。

(3) パターンの配置

文字の配置は、パターンがどのように繰り返し配置されているのかを定量化する。特徴の検出において検出された、パターンがどのように並んでいるのかの情報を用いて定量化を図3のように行う。選択したパターンより下にある検出パターンのうち、最も右上に位置するパターンとなす直線を求め、その直線がY軸に対してどのような角度になっているのかを計算する。

(4) 錠剤の大きさ

錠剤の大きさは、錠剤部分の面積の割合をを求めることで定量化する。錠剤だけに分離された画像を用いて、パッケージ画像全体に対する画素数の割合を求める。

(5) 錠剤の数

錠剤の数は、パッケージにいくつ錠剤があるのかを数え定量化する。錠剤は図4の

ように縦横に等間隔に並んでいる。そこで、錠剤だけに分離された画像から錠剤部分の縦横のヒストグラムを作成し、ヒストグラムの山の数からいくつ並んでいるのかを求め、そして掛け算をすることで錠剤の数が求められる。

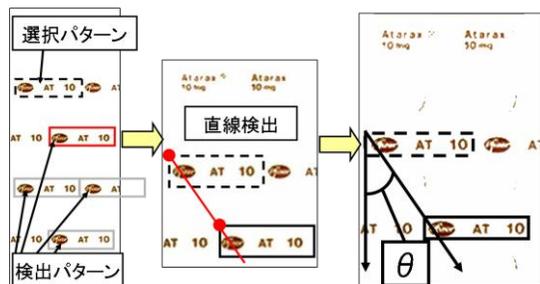


図 3 パターンの配置の定量化

Figure 3 Quantitation of the feature of the arrangement of patterns.

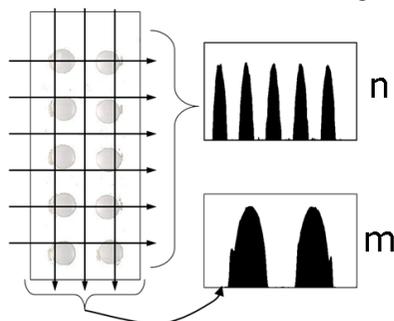


図 4 錠剤の数の定量化

Figure 4 Quantitation of the feature of the number of tablets.

3. 類似性の可視化システム

2.1 節で述べた七つの特徴について類似性を可視化する。本研究で用いる可視化データは多次元のデータであり、そのようなデータに対する既存の可視化手法としては、Parallel Coordinates[5]や Dimensional Stacking[6]などがある。本研究におけるデータの可視化には、人間が見たときに関係性が直感的にわかりやすい二次元グラフを用いる。グラフにプロットを行う際、それぞれのパッケージを錠剤のサムネイル画像で配置し、その座標が何の錠剤なのかを一目でわかるようにする。

作成した可視化システムを図 5 に示す。図 5 では 460 種類の錠剤の類似度を可視化した。本システムでは、全体的な類似の傾向を見ることができる overview mode や、特定のパッケージに対する類似度を見ることができる compare mode のより類似度を可視化する。図 5 (a)に示す overview mode では、任意の二つの特徴について散布図としてプロットして類似度を可視化する。プロットされた錠剤の密集具合などを見ることでその特徴に関する類似の傾向を調べることができる。また、注目したい錠剤や、新たに作成した錠剤パッケージを組み込み、その類似性を表示することで既存の錠剤の中でどのような傾向になっているかを見ることができる。図 5 (b)に示す compare mode では、特定の錠剤に対するその他の錠剤の類似度を可視化することができる。compare mode では原点と比較対象の錠剤をプロットし、その錠剤に対する類似度を原点からの距離によって可視化する。原点に近いほど類似した錠剤となる。X 軸と Y 軸には一つだけでなく複数の特徴を割り当てることができ、複数の特徴に対する類似度を可視化することができる。複数の特徴を割り当て際の重みは、現在では均等に与えている。

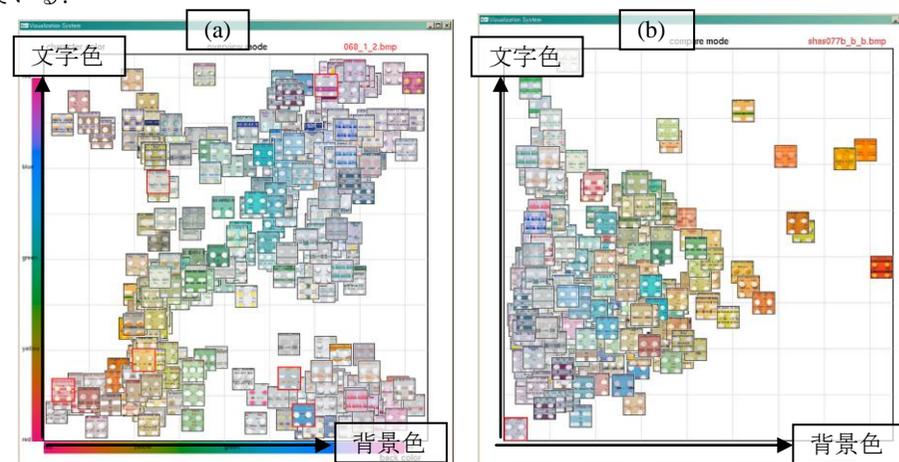


図 5 医薬品の類似度の可視化システム

Figure 5 The visualization system of the similarity of medicines.

4. 文字パターンの類似度

2 章で扱った特徴の他に、異なる錠剤を比べる場合には印刷される文字やマークなどのパターンが同じかどうかを比べ、ということも考えられる。つまり、異なる二つの錠剤における文字パターンの相違も重要な特徴と考えられる。この文字パターン

の類似度を可視化し、3章で述べたシステムを用いて表示した結果が図6である。図6ではy軸に文字パターンの類似性を割り当てた。可視化結果を見てみると、一番類似しているものでもある程度の距離を持ってプロットされていることがわかる。そこで、いくつかのパッケージ画像に対する類似度の数値を見てみると、図6に示したような類似度となった。類似度を示したパッケージと、原点にプロットされたパッケージの文字画像を図7に示し、比較する。図7(d)のパッケージCのように、原点のパッケージと見た目が大きく異なるものについては類似度が余り高くないが、あまり似ていないということが類似度からも読み取れる。しかし、原点の錠剤と錠剤Aは同じ錠剤の容量が異なるものであり、見た目では文字が類似しているように見えるが、類似度を計算すると値がそれほど高くないという結果になった。このように類似したパッケージについて類似度があまり高くない原因を調べるため、図8に示す文字パターンの類似度の計算過程を可視化するシステムを作成した。このシステムを用いて、文字パターンの類似度の計算手法が妥当であるかの確認を行った。

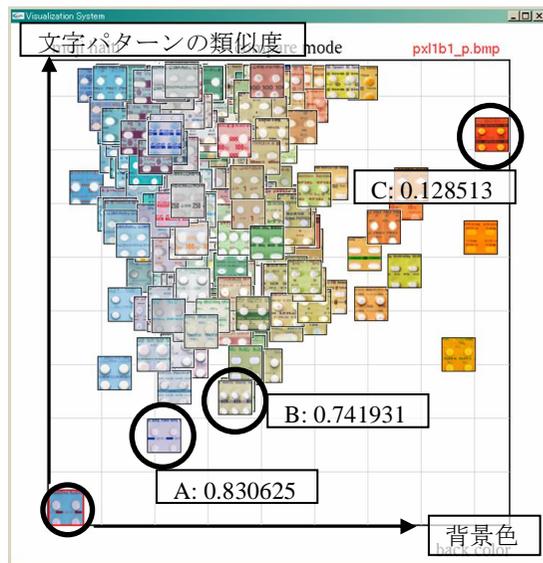


図6 文字パターンの類似度の可視化結果
Figure 6 The result of visualizing the similarity of character patterns.

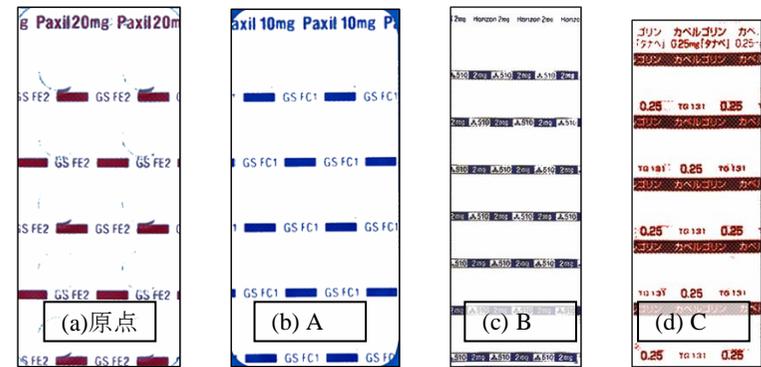


図7 図6で示した錠剤の文字画像
Figure 7 Character images of tablets in figure 6.



図8 文字パターンの計算過程を可視化するシステム
Figure 8 The system which visualizes calculating process of character patterns.

4.1 計算手法

文字パターンの類似性の計算手法について述べる。以下の手法では、文字のパターン画像を2値化したものを用いる。

(1) 提案手法1- 安本らの手法

安本らが提案した文字パターンの類似度の計算手法[7]を説明する。この手法の流れを図8に示す。まず、図9(a)のように2値化したパターン画像の縦横のヒストグラムを得る。次に、図9(b)のようにそれぞれのヒストグラムを重ね合わせ、ヒストグラム

の差が最小となることを探索する。さらに、図9(c)のようにヒストグラムの差が最小の値に近い場所をズレの候補として保存する。そして、図9(d)のように先ほど計算したズレの候補を元に2値画像を重ね合わせ、一致した画素数が最大となるものを探す。そのときの一致画素数の画像の大きさに対する割合を類似度とする。

(2) 提案手法2 - 安本らの手法の改良

提案手法2では、ヒストグラムを比較しズレを求める際、提案手法1のように候補を探すのではなく、ヒストグラムが最も一致するときのズレを計算する。そのズレに対して重ね合わせを行い、一致画素数を求める。

(3) 提案手法3 - 2次元的な比較

2次元的な比較による手法を説明する。提案手法3では、テンプレートマッチングのように2値画像を直接重ね合わせて比較する。最も文字部分の画素が一致する位置を探索し、そのときの一致画素数の割合を類似度とする。

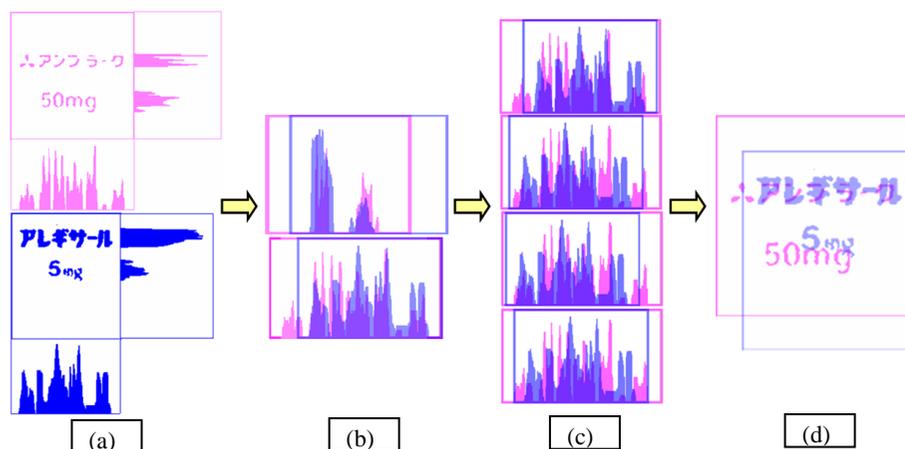


図9 提案手法1の流れ

Figure 9 The flow of the proposal method 1.

4.2 提案手法の計算結果

図8で示したシステムを用いて、4.1節で述べた手法の計算過程について見る。比較のため手動で位置合わせを行ったものも示す。図10に各手法の計算に用いた錠剤のパターン画像の例を、図11にそれらの計算結果を示す。

提案手法1では、図11(a)のように手動の場合とほぼ同じ場所を探せるときがあるが、図11(b)のようにまったく合わないような場合もある。

提案手法2では、図11(c)のように計算結果が提案手法1に比べて少しずれる結果となった。提案手法1や提案手法2では、ヒストグラムの比較によりずらす量を決定しているため、画像の情報を考慮していない。そのため、必ずしも手動で探した場合のように最も一致するズレ量を見つけることができない。

提案手法3では、手動で最も一致する場所を探したときと同じ場所を計算することができた。提案手法1, 2と異なり、画像を直接比較するため最も一致する場所を探し出すことができたと考えられる。

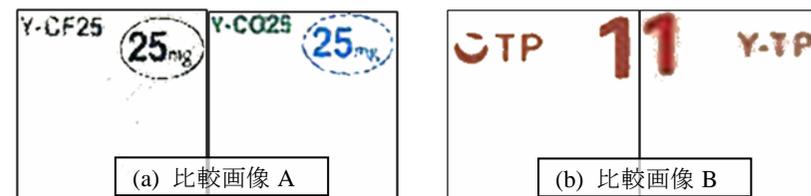


図10 計算に用いた錠剤のパターン画像

Figure 10 The pattern images of tablets used in calculation.

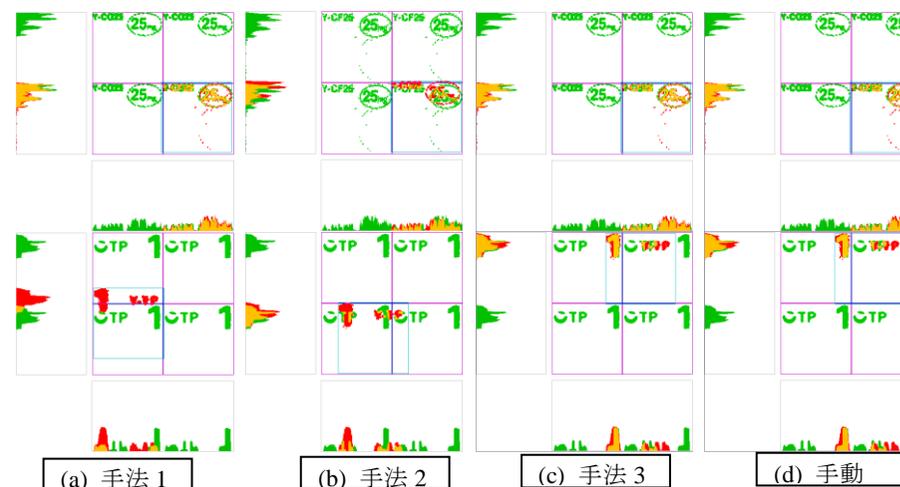


図11 各手法の計算結果

Figure 11 The result of calculation by each method.

4.3 計算手法の改善案

提案手法3では最も一致する場所を探すことができるが、他の手法に比べ計算に時間がかかってしまう。実際に計算時間を計ったところ、縦横ともに152ピクセルである2枚の画像を1回比較するのにかかる時間は平均1.9秒となった。類似度の可視化システムでは現在460枚の画像を扱っており、特定のパッケージについて文字パターンの類似度を計算するには、1回の計算に15分弱かかってしまう。そのため、計算時間を減らすような手法が求められる。そこで、この問題の改善案として、提案手法1を利用することを考える。提案手法1では、ヒストグラムが一致すると思われる候補をいくつか探索する。これらの候補を利用して、2次元で探索する範囲を限定することができれば、計算時間の短縮ができると考えられる。

また、現在は2値画像により計算を行っているが、2値画像ではなくグレースケール画像やカラー画像を用いる手法が考えられる。このとき、2値画像とは異なり、がその有り無しで一致した画素数を求めることができない。そのため、グレースケール画像やカラー画像に対応した類似度の計算手法を考える必要がある。例えば、画素が一致した数を計算する場合、各画素の輝度値により重みをつけたものを求めることで、類似度を計算できると考えられる。

5. おわりに

本研究では、医薬品パッケージの類似性を検出し、可視化するシステムを提案した。また、新たな特徴として文字パターンの類似度についての計算手法を提案した。

今後の課題として、文字パターンの類似度の計算時間の削減、計算に用いる文字画像の種類の検討、現在計算に用いている文字画像の取り直しが上げられる。可視化システムに関する課題としては、銀色の錠剤に対する処理、フィルタリング機能の実装などが上げられる。

参考文献

- 1) 医療事故防止に関する検討会: 医療事故防止対策ガイドライン, 大阪府 (2000).
- 2) J.R. Smith, Shih-Fu Chang: Single Color Extraction and Image Query, 1995 International Conference on Image Processing (ICIP'95) - Volume 3, pp.3528-3531 (1995).
- 3) Smith, J.R., Shih-Fu Chang: Local Color and Texture Extraction and Spatial Query, IEEE International Conference on Image Processing, 1996. Proceedings., Volume 3, 16-19 Sept., pp.1011-1014 (1996).
- 4) Liu, Y., Collins, R., and Tsin, Y: A Computational Model for Periodic Pattern Perception Based on Frieze and Wallpaper Groups, IEEE Trans. PAMI 26, 3, pp.354-371 (2004).
- 5) Inselberg, A., and Dimsdale, B.: Parallel coordinates: a tool for visualizing multi-dimensional geometry, Proc. of Visualization '90, pp.361-378, (1990).

6) J.LeBlanc, M.O.Ward and N.Wittles: Exploring N-dimensional Databases, In Proc. Visualization '90, pp.230-239, (1990).

7) 安本聖理奈, 古谷雅理, 宮村浩子, 斎藤隆文: 医薬品パッケージの類似性の検出と可視化, 第71回情報処理学会全国大会論文誌, 3Z-5 (2009).