

歯科情報による個人検索支援システムの開発と評価

安田 大誠^{1,a)} 吉野 孝^{1,b)} 玉川 裕夫^{2,c)}

概要：2011 年 3 月 11 日に発生した東日本大震災の経験から、身元不明のご遺体に対する身元確認の手段のうち、歯型鑑定による方法の有用性が報告されている。具体的には、津波などの住宅が消失するほどの被害を受けた場合においても、掛かりつけの歯科医院に本人の情報が存在すれば、身元確認をすることが可能である。さらに、近年発生予想がされている南海トラフ巨大地震では、東日本大震災を上回る被害想定がなされている。そのため、歯型鑑定がより効果的に用いられる可能性がある。また、厚生労働省の「歯科情報の標準化」事業により、各歯科医療機関で蓄積される歯科情報に対し、標準的なデータ形式を定義する取り組みが行われている。そこで我々は、厚生労働省の事業により社会活用できる段階となった歯科診療情報を用い、身元確認を支援するシステムの開発を行っている。システムの特徴として、歯が持つ階層構造に着目し、検索時の工夫としている。階層構造にすることで、災害時の歯の損傷などに対応が可能となっている。本稿では、現時点で 3 層まで実装したシステムを用いた評価実験により、以下の 2 点を明らかにした。(1) 3 層までの階層構造は、歯 4 本分までの表記揺れに対して、精度よく検索することが可能である。(2) 3 層まで階層を下げて検索するにつれ、検索精度を高めることができる。

キーワード：東日本大震災、南海トラフ巨大地震、個人識別、身元確認、歯科診療情報の標準化、検索システム、階層構造

1. はじめに

2011 年 3 月 11 日に発生した東日本大震災では、ご遺体の個人識別（身元確認）が、主に次の 4 通りの方法で行われた。「身体的特徴や所持品による個人識別」「指紋・掌紋による個人識別」「DNA 型による個人識別」「歯科的特徴による個人識別」である。その中でも、「歯科的特徴による個人識別」が有効であった。具体的には、東日本大震災で亡くなった岩手・宮城・福島 3 県の犠牲者の身元確認のうち、歯型鑑定によるものが DNA 鑑定の約 7 倍であり、宮城県単体でみると約 10 倍となっている [1][2]。

しかし、震災当時、警察でさえも、多数のご遺体の歯科的个人識別について十分な経験を有していなかった。そのため、青木らは歯科医師会と合同で、歯科情報に基づく迅速な身元確認のワークフローを構築した [3]。

近年発生予想がされている南海トラフ巨大地震では、東日本大震災を上回る被害想定がなされている。

また、次章で述べる「歯科診療情報の標準化」事業が行われ、今や歯科情報が社会活用できる段階となっている。

そこで本稿では、一定の地域内で蓄積された歯科情報を検索する前段階として、診療所内の外部記憶装置に蓄積された情報のテキスト検索を想定したシステムを提案する。

2. 歯科診療情報の標準化

震災の教訓から、将来の大規模災害の身元確認に備えるために、「歯科診療情報の標準化」事業が厚生労働省に設置され、実証事業が平成 25 年から実施されている [4]。この事業の目的は、各歯科医療機関で蓄積される歯科情報に対し、標準的なデータ形式を定義し、歯科情報を社会的に活用できる仕組みの実現である。

平成 25 年度事業では、歯牙特徴に基づいた「標準プロファイル 26 項目」[5] が策定され、身元確認において極めて高精度に絞り込みが可能であることが実証された。

平成 26 年度事業では、「口腔状態の標準データセット」[6] が策定された。一般に、歯科医療機関で用いられている診療報酬明細書（レセプト）作成用コンピュータ（以下レセコンと略）及び電子カルテの内部データについては、ベンダごとに仕様が全く異なっている。それらのデータの詳細度（粒度）については、統一規格が存在しない。「口腔状態標準データセット」は、歯が持つ情報の粒度に応じて

¹ 和歌山大学
Wakayama University, Wakayama 640-8510, Japan

² 大阪大学歯学部付属病院
Osaka University Dental Hospital, Osaka 565-0871, Japan

a) yasuda.taisei@g.wakayama-u.jp

b) yoshino@sys.wakayama-u.ac.jp

c) tamagawa@dent.osaka-u.ac.jp

階層的に構造化した特徴記述子を用いている。これによって、レセコンや電子カルテのベンダが、それぞれの製品に応じた情報の粒度でデータを抽出・保存し、活用することができるようになった。

平成 27 年度事業では、「口腔状態の標準データセット」などをもとに、レセコン等からデジタルデータを出力するための仕様書「口腔診査情報標準コード仕様」[7] が策定された。「口腔診査情報標準コード仕様」の策定によって、各歯科医療機関で蓄積される歯科情報を同じ形式で管理できるようになるため、歯科医療機関をまたいだデータ活用が可能となった。

平成 28 年度事業では、活用可能となった歯科情報の保存方法・利活用方法の検討が行われ、CSV 形式のデータを標準として利用することとなった。

平成 29 年度からは事業名を「歯科情報の利活用及び標準化普及」と改め、2 地域ほどのモデル地区（地域医療ネットワーク）で「口腔診査情報標準コード」に準拠した電子カルテ等の実証が行われた。一定の地域内で共通形式のデータを蓄積し、それらのデータに対する歯式ビューアが開発されるなどした。

3. 関連研究

3.1 歯科情報を用いた身元確認に関する研究

青木らは、生前カルテ情報とご遺体情報を照合する専用ソフトウェア Dental Finder を開発・運用した [3]。Dental Finder では、各々の歯の状態を 1～5 の 5 分類符号で表現し、口腔内の所見を 32 桁の数字列に置き換える。この数字列の類似度によって、生前・死後情報の検索を行う。

菊月らは、「有」「無」「不明」の 3 種類に 歯科所見を簡素化して絞り込みを行う「36（サブブロック）検索」を開発した [8]。36 検索は、上下左右の犬歯と第一大臼歯の計 8 本を、「どの歯があって、どの歯がないか」という法歯学の基本に則って設計されている。

どちらの研究においても、岩手県、宮城県、福島県などの地域や、その地域内ごとに、歯科診療情報の互換性がないという問題を抱えており、データ連携に時間を要した。また、データ連携を行わない場合、身元確認をするには、歯科医療機関ごとの診療情報の形式に適したシステムの開発が必要となる。本研究とは、互換性のある標準的なデータ形式に基づいて設計されたという点で異なる。

3.2 標準的なデータ形式に基づいて身元確認が行われた事例

海外では、大規模災害によって多数遺体が発生した場合の身元確認に対して、DVI (Disaster Victim Identification) という専門用語が用いられている。2004 年のインド洋津

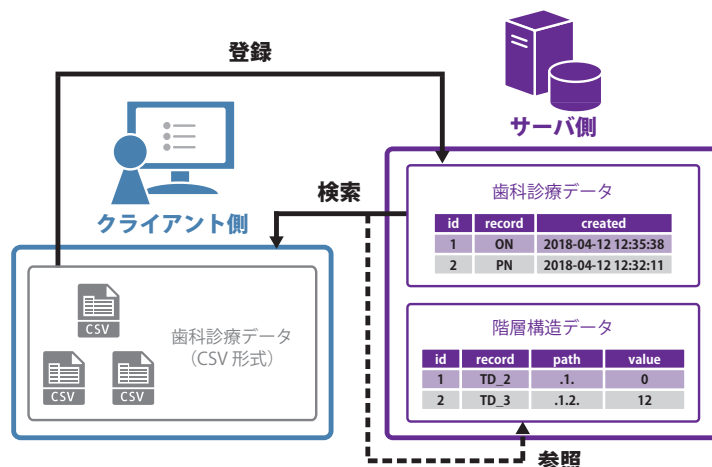


図 1 システム構成

波被災後のタイでは、INTERPOL *1 の協力を受けて DVI が国家事業として行われた [9][10]。このとき使用されたソフトウェアでは、各々の歯の状態に関して、歯の状態、補綴物、X 線などの 6 つのカテゴリで表現する。このカテゴリは、INTERPOL が作成した、行方不明者や死者を特定するための書式用紙の仕様に従っている。ゆえに、互換性のある標準的なデータ形式に基づいて設計されたという点では同様である。しかし、本研究とは各々の歯の状態を、階層的に構造化した特徴記述子を用いて表現している点で異なる。

4. 歯科情報照合システム

4.1 システム構成

図 1 にシステムの構成を示す。本システムは、「口腔診査情報標準コード仕様」に基づいた CSV 形式の歯科診療データに対応した、Web ベースのシステムである。なお、本稿における「CSV 形式の歯科診療データ」とは、「診療所内の外部記憶装置に蓄積された情報」と同義である。本システムは、ユーザが操作するクライアント側と、歯科診療データ及び階層構造データを管理するサーバ側から構成される。クライアント側ではサーバ側に対し、「CSV 形式のファイルの登録」「登録されているファイルの検索」を実行することができる。サーバ側では、RDB (Relational Data Base) の MySQL *2 で管理を行う。CSV 形式のファイルには 1 ファイルにつき、患者の 1 回の診察に関する情報が記録されている。このファイルをサーバ側に登録するときには、データの粒度に応じ、正規化した MySQL のテーブル形式で保存する。また、歯科診療データとは別に、歯の状態を階層的に構造化した特徴記述子を管理者が登録している。このシステム構成により、「診療所内の外部記憶装置に蓄積された情報のテキスト検索を想定」に関して、「サー

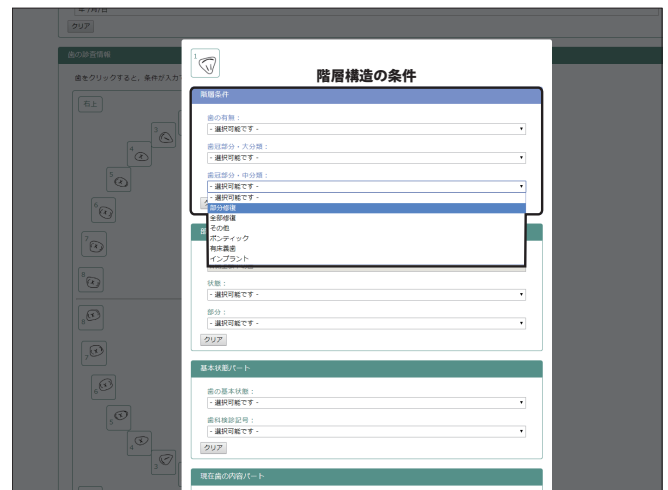
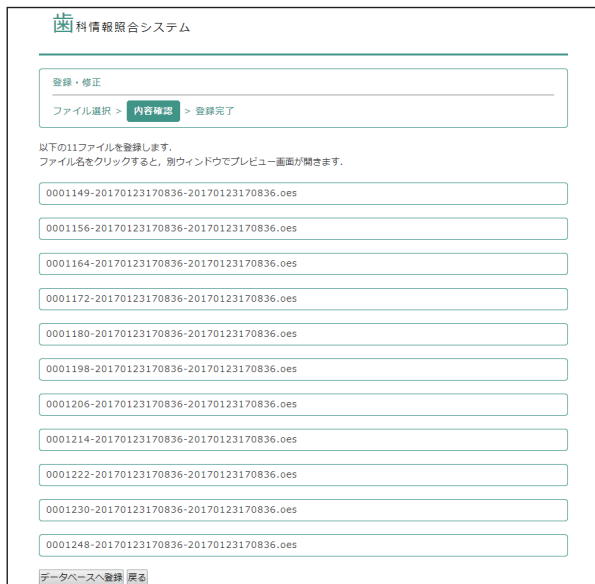
*1 International Criminal Police Organization (国際刑事警察機構): <https://www.interpol.int/>

*2 MySQL: <https://www.mysql.com/jp/>

(1) ファイル選択画面



(2) 登録ファイル確認画面



フォーム画面」で入力し、「検索結果画面」で確認するという順序で操作する。

(1) 検索フォーム画面

基本的には、図 2 と同じ画面構成となっている。図 5 に、歯のアイコンをクリックしたときモーダル表示される、検索フォーム画面を記す。青色の部分は、後述する階層構造データを用いた検索のフォーム、緑色の部分は、診療情報の検索のフォームとなっている。この画面では、歯に関する情報を入力することが可能であるが、4.2.1 項 (4) に記したように、現在歯と欠損歯の内容は排他関係であり、どちらか一方の内容しか入力できない。

(2) 検索結果画面

図 6 に検索結果画面を示す。この画面は、主に (a), (b), (c) の 3 箇所から構成される。

(a) 検索結果 (全条件: AND)

(a) には、入力した条件のうち、共通して当てはまる結果のみが表示される。このとき、全条件をひとまとまりに捉えるのではなく、次のロジックを用いて歯の位置を考慮し検索を行う。

$$A_n = B_{n1} \cap B_{n2} \cap \dots B_{nm}$$

$$X = A_1 \cap A_2 \cap \dots A_n$$

$$\left(\begin{array}{l} X: \text{検索結果,} \\ A_n: n \text{ 番目の歯に関する入力条件,} \\ B_{nm}: n \text{ 番目の歯に関する } m \text{ 番目の} \\ \text{入力条件} \end{array} \right)$$

ここで表示される 1 ファイルは、1 患者を示している。また、ファイル名をクリックすることで 4.2.1 項の基本情報のプレビュー画面が表示される。

(b) 階層構造条件の再検索

(b) は、検索条件に階層構造の条件が含まれていた場合、表示される。階層構造の条件のうち、最

録時の実行時間を考慮し、50 ファイルまでとしている。

(c) ファイルプレビュー (CSV 形式)

(b) でファイル選択が完了すると、ファイル名とともに CSV 形式でプレビューが行われる。

(2) 登録ファイル確認画面

この画面ではファイル登録の最終確認を行う。(1) で選択されたファイル名がリスト表示され、クリックすることで 4.2.1 項の基本情報のプレビュー画面が表示される。

4.2.3 検索機能

本機能が本システムの根幹をなす部分である。「検索

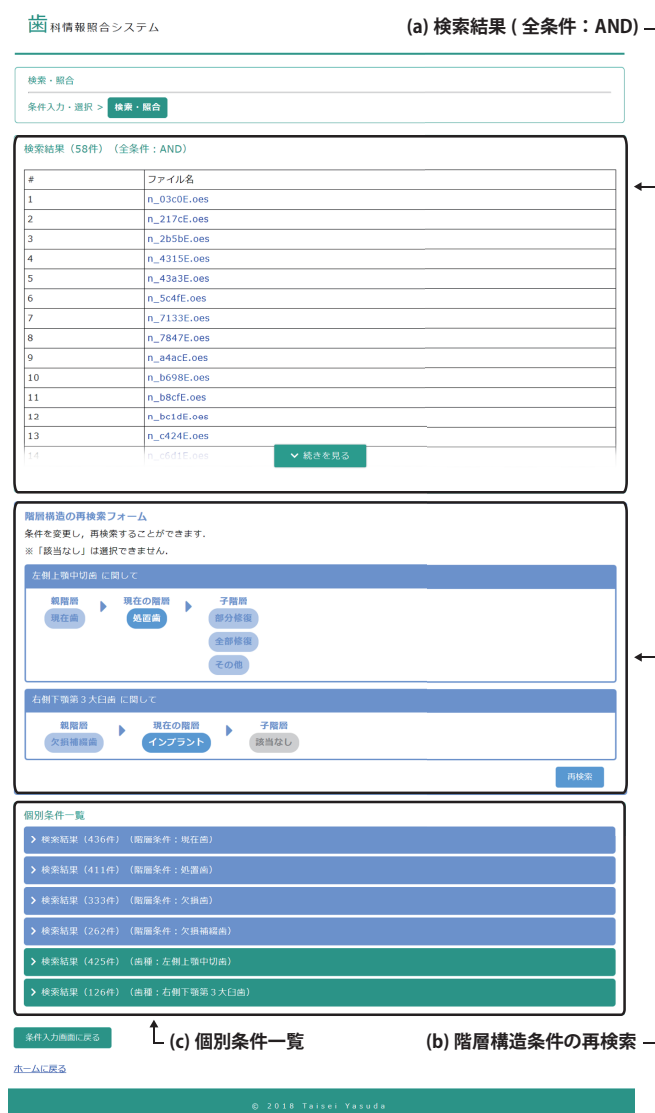


図 6 検索結果画面例

下層に該当するものが、「現階層」として表示される。1つ上と下の概念はそれぞれ「親階層」「子階層」として表示され、クリックすることでその条件も含めて再検索することができる。

(c) 個別条件一覧

(c) は、検索条件をそれぞれ個別のものと扱った場合の検索結果を表示する箇所である。クリックすることで検索結果リストが開き、更にファイル名をクリックすることで 4.2.1 項の基本情報のプレビュー画面が表示される。

(b) や (c) を用いることで、結果の更なる絞り込みなどに活用することができる。

(3) 階層構造の条件検索について

図 7 に歯の階層構造例を示す。この階層構造は、「口腔状態の標準データセット」[6] を参考に、「口腔診査情報標準コード仕様」[7] 内の歯の診査情報レコードに対応させたものとしている。なお、図 7 中の T(ア

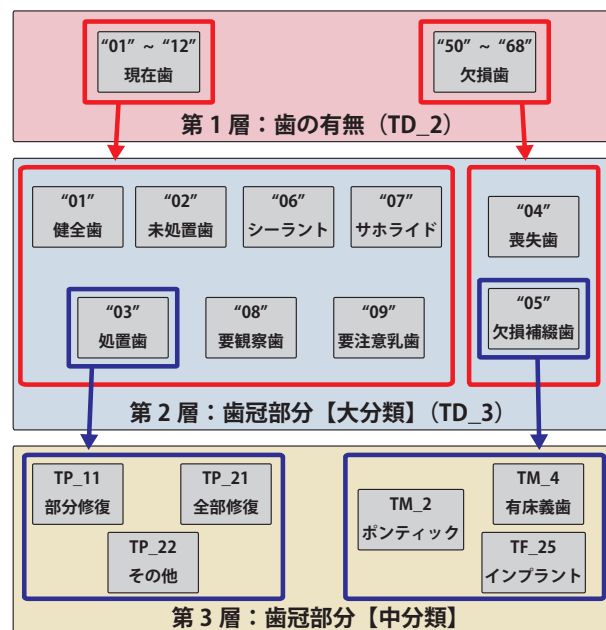


図 7 歯の階層構造例

ルファベット)_(数字) は、そのレコード内の情報である。第1層は、歯そのものがある・ないの2値、第2層は、歯冠部分（口の中に出ているところ）の大分類、第3層は、歯冠部分（口の中に出ているところ）の中分類、で判別する。

本システムでは、階層構造のデータ構造を保存するのに、経路列挙モデルを用いた。経路列挙モデルには、各ノードまでの絶対パスをデータとして保存している、という特徴があるため、親子関係の経路探索のクエリが簡単になる。

表 1 に例を示す。id には、各データをユニークにするための番号を振り分ける。record には、図 7 中の T(アルファベット)_(数字) を1つずつ格納する。name には、データを表す歯の状態を格納する。path には、第1層をルートとした、そのデータまでの id の経路を格納する。value には、record が同じだった場合の識別子として用いられる、数値を格納する。

例として、id が 14 で、record が TP_21 の「全部修復」について説明する。図 7 を参考に「全部修復」の親をルートまでたどると、「処置歯」「現在歯」となる。ゆえに、「全部修復」の path は「処置歯」の id:6 と、「現在歯」の id:2 を用いて「1.2.6.14」と表せる。

このモデルは、階層条件を用いた検索と、検索結果画面の (b) 階層構造条件の再検索で、必要があった場合に参照される。

5. 評価実験

5.1 本実験のデータ構成

本実験では、2017 年に新潟県と静岡県で収集された 436

表 1 経路列挙モデルの例

id	record	name	path	value
1	NULL	NULL	.1.	NULL
2	TD_2	現在歯	.1.2.	12
3	TD_2	欠損歯	.1.3.	50
4	TD_3	健全歯	.1.2.4.	1
5	TD_3	未処置歯	.1.2.5.	2
6	TD_3	処置歯	.1.2.6.	3
7	TD_3	シーラント	.1.2.7.	6
8	TD_3	サホライド	.1.2.8.	7
9	TD_3	要観察歯	.1.2.9.	8
10	TD_3	要注意乳歯	.1.2.10.	9
11	TD_3	喪失歯	.1.3.11.	4
12	TD_3	欠損補綴歯	.1.3.12.	5
13	TP_11	部分修復	.1.2.6.13.	NULL
14	TP_21	全部修復	.1.2.6.14.	NULL
15	TP_22	その他	.1.2.6.15.	NULL
16	TM_2	ボンティック	.1.3.12.16.	NULL
17	TM_4	有床義歯	.1.3.12.17.	NULL
18	TF_25	インプラント	.1.3.12.18.	NULL

人分の CSV ファイル^{*3}（以下、実験ファイル）を検索データとして使用した。

歯科情報を用いた個人検索の場合、被検索側の「生前データ」と、検索側の「死後データ」が必要となる。これら 2 つのデータを照合することによって、個人の検索が初めて可能となる。一般的に「生前データ」は、歯科医療機関で管理されている。本実験では、実験ファイルをファイル登録機能でサーバに登録することで、同様の環境を用意した。「死後データ」は災害後に、ご遺体から法歯学者が得るデータである。本実験では、実験ファイルに外乱を適用することによって、仮想の「死後データ」を用意した。

5.2 検証項目

本実験では、階層構造によるシステムの検索性能を評価するために、以下の項目について検証を行った。

- (1) 階層構造は、どのくらいの表記揺れまで許容し、検索できるか
- (2) 階層構造での絞り込みのしやすさはどのくらいか

5.3 検証概要

以下に、検証項目の具体的な内容について述べる。

5.3.1 「階層構造は、どのくらいの表記揺れまで許容し、検索できるか」について

本システムは、階層構造データにより、検索現場での表記に揺れが発生しても、それを許容し検索することが可能である。ゆえに、本項目の目的は、表記揺れをデータ上の外乱とみなして検索を行い、検索されにくくなる所との閾値を見つけることである。外乱を適用した歯の本数によって、検索精度がどのように変化するかを確認し、閾値について考察する。本実験における検索精度とは、「絞り込み」を意味する。そこで、本実験では検索結果数を評価指標とした。これは、目的のファイル以外の結果が多くなれば絞り込めていない、結果が少なくなれば絞り込めている、と判断が可能なのである。

5.3.2 「階層構造での絞り込みやすさはどのくらいか」について

本項目での目的は、各層における検索精度を比較することである。本項目においても、検索結果が少ないほど絞り込みができていないと判断可能なため、検索結果数に着目し、評価指標として用いた。手順としては、最上位の第 1 層から最下位の第 3 層まで再検索機能を用いて、検索を行う。このとき、階層ごとに、どう検索結果数が変化するかを確認する。

5.4 外乱の適用

データの外乱は、歯を単位として、ランダムに選出した歯に適用する。本実験で適用する外乱は、以下の 2 種類である。

（タイプ A）歯の状態が不完全：階層の粒度を下げる

（タイプ B）歯の死後脱落：その歯に関する情報を無くす

本実験においては、全ファイルの 5% にあたる、22 件の実験ファイルを 9 回サンプリングし、外乱を適用した。22 件のうち、半分の 11 件は固定で、もう半分の 11 件は重複無しのランダムに実験ファイルをサンプリングしている。半分の固定にしたのは、比較を行えるようにするためである。そして、サンプリングした 22 件、9 セットの実験ファイルに対して、0 本分（選出しない）から順に 8 本分まで、情報を持つ歯をランダムに選出する。最後に、選出した歯に対して、2 種類の外乱のどちらかをランダムに適用する。これらは全て Python^{*4} で実装した。

6. 実験結果と考察

6.1 「階層構造は、どのくらいの表記揺れまで許容し、検索できるか」の評価

階層数と、歯に適用した外乱の数ごとに検索を行った。このとき、適用した外乱数以外は同条件である。検索結果数をもとに、適用した外乱数の条件ごとの母集団正規分布

^{*3} 平成 29 年度 歯科情報の利活用及び標準化普及事業 (2) 診療情報共有を目指したモデル地区展開、厚生労働省：https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/shiryoi_3.pdf

^{*4} Python：<https://www.python.org/>

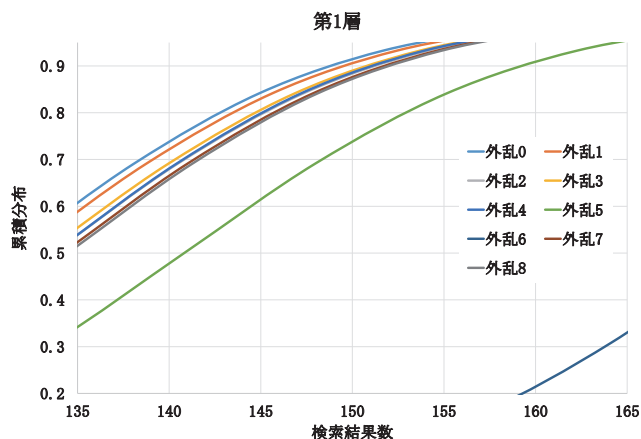


図 8 第 1 層の検索結果

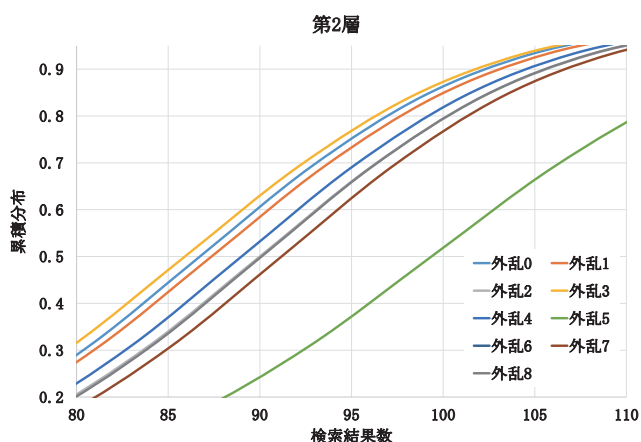


図 9 第 2 層の検索結果

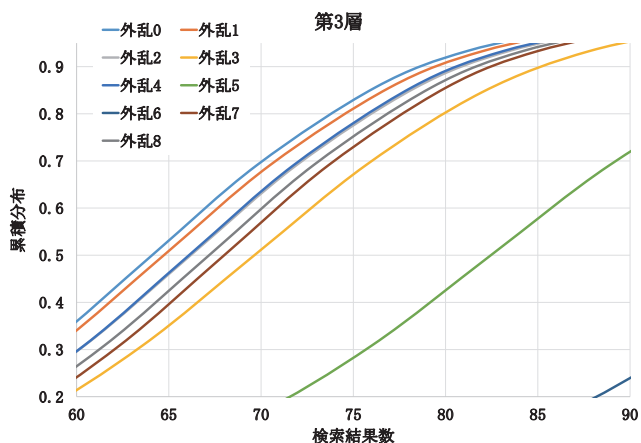


図 10 第 3 層の検索結果

を推定した。母平均推定と、母分散推定にはそれぞれ 95% 信頼区間の t 分布と、 χ^2 分布を用いた。

図 8 に第 1 層の検索結果、図 9 に第 2 層の検索結果、図 10 に第 3 層の検索結果を示す。これらのグラフは横軸に検索結果数、縦軸に累積分布をとっているため、グラフが左上に位置するほどより検索結果数が少なくなる。従って、グラフが左上に位置するほど精度が高くなる。精度が高い順

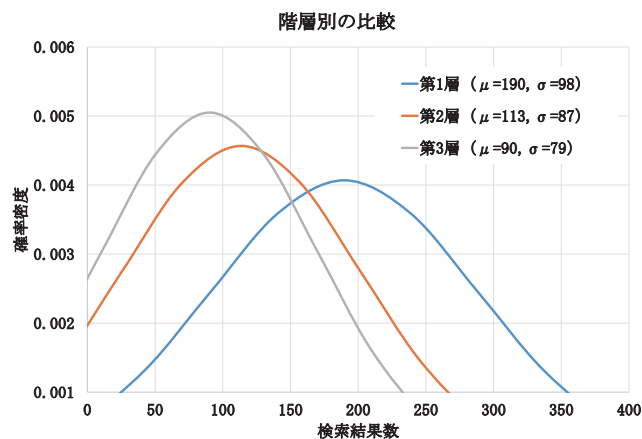


図 11 階層別の検索結果

に外乱数が、図 8 では $0 > 1 > 3 > 4 > 2 > 7 > 8 > 5 > 6$ 、図 9 では $3 > 0 > 1 > 4 > 2 > 8 > 7 > 5 > 6$ 、図 10 では $0 > 1 > 4 > 2 > 8 > 7 > 3 > 5 > 6$ 、となっている。このことから、比較的外乱が多いと、検索精度が低くなることが分かった。また、検証項目である閾値については、外乱数 4 とした。理由としては、図 8、図 9、図 10 の検索結果から、外乱数 5 になると急激に検索精度が下がるためである。しかし、外乱数が 7、8 になると、検索精度が再び上がってしまっている。これは、外乱を適用しすぎたため、元の状態から大きく離れ、検索の条件にユニークさが出てしまい、逆に結果を絞り込めてしまったと考えられる。図 9 の第 2 層の検索結果で、外乱数 3 の方が、外乱数 0 (適用していない) より検索精度が高く出ているのも、同様の理由であると考えられる。

6.2 「階層構造での絞り込みやすさはどのくらいか」の評価

実験結果 1 の考察から、適用した外乱数が 4 より多い場合は除外したうえで、各階層の検索結果数を集計し、検索結果数の母集団正規分布の推定を行った。集計したのは、外乱数が 0 から 4 まで適用したときの、重複無しランダム 11 件、数にして 55 件のファイルに対する、各階層の検索結果数である。母平均推定にはそれぞれ 95% 信頼区間の正規分布を用いた。

図 11 に階層別の検索結果を示す。第 1 層の検索結果数は μ が 190、 σ が 98 の青色の分布、第 2 層の検索結果数は μ が 113、 σ が 87 の橙色の分布、第 3 層の検索結果数は μ が 90、 σ が 79 の灰色の分布、である。分布の平均値は、第 1 層、第 2 層、第 3 層、というように検索結果数が少なくなり、第 1 層から第 3 層までは半分以下となっている。さらに、平均値が下がるだけでなく、標準偏差も小さくなることから、分布のばらつきが小さく、検索の精度自体も向上している。これらのことから、階層を下げて検索していくにつれて、検索精度が高くなることが分かった。

7. おわりに

本稿では、歯科情報を用いた身元確認支援システムの機能の説明と、現時点で3階層までの階層構造によるシステムの検索性能の評価実験について述べた。性能評価実験の結果から、以下の2点を明らかにした。

- (1) 3層までの階層構造は、歯4本分までの表記揺れに対して、精度よく検索することが可能である。
- (2) 3層まで階層を下げて検索するにつれ、検索精度を高めることができる。

実験結果から、階層数を増やすことでシステム性能の更なる向上余地があると考えられるため、今後の予定として、より深い階層の検討と、マッチングアルゴリズムの模索を行っていく。

参考文献

- [1] 平成 24 年警察白書 統計資料：特-5 東日本大震災における遺体の身元確認状況について（平成 24 年 5 月 11 日現在）, <https://www.npa.go.jp/hakusyo/h24/toukei/00/0-05.xls> (参照：2018-07-18).
- [2] 日本経済新聞：震災犠牲者の身元確認、DNAより歯型が有効 (2016-03-13), <https://www.nikkei.com/article/DGXLZ098375320S6A310C1000000/> (参照：2018-07-18).
- [3] 青木孝文, 伊藤康一, 青山章一郎：災害犠牲者の身元確認と ICT, IEICE Fundamentals Review, 9 (2), pp.119-130, 2015.
- [4] 厚生労働省：歯科情報の標準化について, <https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000205867.html> (参照：2018-07-18).
- [5] 厚生労働省：第 5 回歯科診療情報の標準化に関する検討会 資料 1, <https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000076199.pdf> (参照：2018-07-18).
- [6] 厚生労働省：口腔状態の標準データセット, <https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000117126.pdf> (参照：2018-07-18).
- [7] 厚生労働省：口腔診査情報標準コード仕様 ver.1.0, <https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000155163.pdf> (参照：2018-07-18).
- [8] 菊月圭吾, 狩野敦史, 小野博之, 白石秀幸, 黒澤正雄, 熊谷哲也, 斎藤雅人, 熊谷章子, 藤村朗, 出羽厚二：東日本大震災における身元確認作業, 岩手医科大学歯学雑誌, 37 巻 2 号, pp.74-84, 2012.
- [9] M. Petju, A. Suteerayongprasert, R. Thongpud, K. Hassisiri : Importance of dental records for victim identification following the Indian Ocean tsunami disaster in Thailand, Public Health 121, pp.251-257, 2007.
- [10] L. Andersen Torpet : DVI System International: Software Assisting in the Thai Tsunami Victim Identification Process, The Journal of Forensic Odonto-Stomatology, Vol.23 No.1, pp.19-25, 2005.