

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 23 日現在

機関番号：21602

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700094

研究課題名(和文) 情報地球儀：球面クラスタリングによるデータの関係の可視化

研究課題名(英文) Information Sphere: A Spherical Visualization for Data Relationship

研究代表者

矢口 勇一 (Yaguchi, Yuichi)

会津大学・コンピュータ理工学部・准教授

研究者番号：00609109

交付決定額(研究期間全体)：(直接経費) 3,300,000円、(間接経費) 990,000円

研究成果の概要(和文)：本研究の目的は、文書や画像を類似度や親近度を基にした球面への再配置問題を解くことである。球面は、ある物体を中心においても、距離、方向などにおいて空間が変わらないため、他のアルゴリズムのように再計算する必要がなくなる。また、地球儀を回すように可視化することができる。成果として、本研究を通して作成されたシステムから、3.11東日本大震災で被害にあわれた方々への復興支援(写真探索システム)や、ビジネスへの転用(コーディネーションマップ、ソフトウェアコンポーネントマップ)などの研究が進んだ。

研究成果の概要(英文)：To solve object re-mapping problem such as documents or images onto spherical surface based on similarity or association, we establish unbiased mapping and visualization from moving origin. Sphere is able to change focusing point because if focusing point will change but density of distance and direction will not change. Also, this sphere is able to visualize directly as globe. From this result, we constructed several system such as lost photo search system for great Tohoku disaster, clothing coordination recommendation, software component map etc.

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：アルゴリズム 可視化 画像 ディレクトリ・情報検索 感性情報学

1. 研究開始当初の背景

文書や画像等のメディアに対して、類似するもののクラスタリング (K-Means, NN, SVM) などは過去に多く研究があり、また、類似度や親近度を行列として与えられた時の自動的なマッピングについても、過去に研究 (数量化 IV 類、自己組織化マップなど) がある。ところで、この自動的なマッピングにおいて、ある物体 A を世界の中心として置く場合、過去研究では、ある物体 A を原点にして、再度アルゴリズムをかける必要があり、また、それでできた空間はユークリッド空間もしくは 2 次元平面であるため、空間の外側、特に端にある点に対しての情報の偏りは中央に対して激しく異なるものとなるため、中心を設定しないで掛けたものと大きな差異が生じる。このため、このような差異が生じにくい空間を定義しマッピングすることは、同時にある物体 A を中心として再定義しても、再計算する必要がなく、再構成された空間中で物体 A を原点もしくは原点に資する位置へと移動してそこから世界を見に行くことができるのではないかというのが本研究の着想である。

2. 研究の目的

本研究の中心となる目的は 2 つある。1 つは、文章や画像などの『類似』をどのように定義するか、また、『類似』や『関係性』を通して用法や傾向のパターンのクラスタリングが可能か。もう 1 つは、この用法や傾向のパターンを用いて、『球面へとマッピングが可能か』と言うことである。ここで、なぜ球面かと言うと、球面は、面として 2 次元の情報を持つが、空間的には 3 次元の情報となるため、2 次元よりも情報が多く、3 次元の情報のコンパクトとなる。また、ある一点からの最遠点は、半径を 1 とする球面であれば必ず 2 である。また、ある点を球面上の中央点として考察したときに、最遠点の 2 までは、球面上の 180 度どの方向に進んだとしても、同じ遠さ、同じ空間密度となる。そのため、背景に記したように、点の中心を再定義しても、再計算する必要なく、単に『地球儀を回すように』情報の位置取りを変えることが可能である。さらに、表現形式が 2 次元であることから、現在のコンピュータのディスプレイ上では、非常に表現しやすいものとなる。また、本手法の元アルゴリズムとなる連想単語空間 (Associated Keyword Space, ASKS) は、親近度に基づく非計量的な多次元尺度構成法であり、非計量的部分を計量的に扱うように置き換えること、また、関連のない部分を 0 と置く (計算しない) と置くことで、文書と画像などの異種間のデータに対しても同時に空間内にマッピングすることが可能であることから、それを用いて、空間上で使用傾向などが類似するもののクラスタリングなどに有効利用することができるものとする。

3. 研究の方法

入力として、文書や画像の類似度や親近度の行列を与える。この類似度や親近度は、マップにしたい時のパラメータとなるが、画像においては相関や画像間の距離、文書などは TF・IDF などの計量的な値や、または文書などの Web リンクなどによる隣接関係など (非計量) も含めることができる。仮にこれを親近度行列としておくと、親近度行列を入力として、3 次元空間へ、近い関係のものは近く、遠い関係のものはより遠くなるように再配置する連想単語空間 (ASKS) アルゴリズムを用いてマッピングする。ASKS 法は多次元尺度構成法と呼ばれる手法の一つで、類似度の高いものを近くに、低いものを遠くに再配置する問題(1)式を解く。

$$J(x_1, x_2, \dots, x_n) = \sum_i \sum_j \{-M_{ij}F(x_j - x_i)\} \rightarrow \max. \quad (1)$$

(1)式において、 x_n は解くべき座標点、 $F(k)$ は k の距離に対するカーネル、 M_{ij} は類似度行列の ij 要素を意味する。

この計算は最急降下法による繰り返し演算を行うが、繰り返しの度に生成される 3 次元点 x に錐制約を加える。つまり、 x に対して極座標系へと変換し、重心からの距離を 1 へと変更することである。これらを繰り返して、球面へとマッピングする。これを球面 ASKS (S-ASKS) と呼ぶ。S-ASKS によって再配置された物体は、可視化ツールによって、Google Maps のように、地球儀を回すような形で閲覧することが可能となる。また、視点の拡大・縮小を行うことで、どれだけ密集した場所であっても、地球儀に寄ってみれば理解できるように、簡単に密集に対する解像度を変えることも可能となった。また、画像などの類似性に対しては、画像を縦横 10 本の柵状にカットした状態でカラーヒストグラムを作成したものから、動的計画法を用いた DP Distance を使用して 2 画像間の類似度を取る事などを行った。これは、画像中に含まれる色と構造を両方保持するためのものである (Fig. 1)。

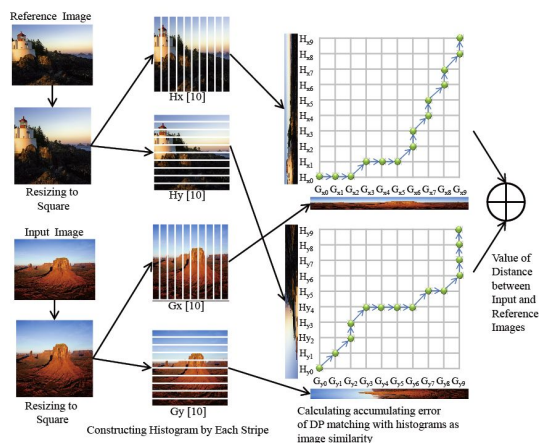


Figure 1. DP 距離に基づく画像の類似度

4. 研究成果

本研究では、成果として、1 件の雑誌論文、11 件の国内外の学会発表を行っており、加えて、現在、1 件の雑誌論文の投稿を目指している。また、S-ASKS による連想空間を用いたいくつかのシステムを考案し、3.11 東日本大震災で被害にあわれた方々への復興支援（写真探索システム）(Fig 2.)や、ビジネスへの転用（コーディネーションマップ (Fig3.)、ソフトウェアコンポーネントマップ）などの研究が進んだ。また、さらに、時系列的な変化も連続して閲覧できるようにする時系列 ASKS, 時系列 S-ASKS も同時に考案され、Twitter 解析などへの応用が期待される。



Figure 2. 福島県新地町で展開した遺失写真探索システム

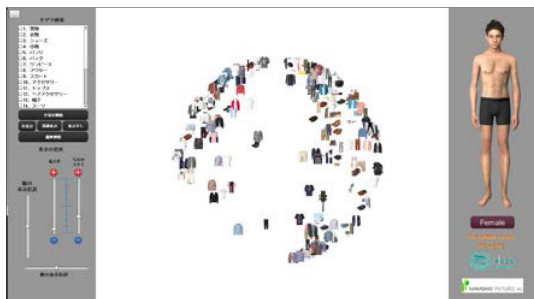


Figure 3. Style Share: ユーザーコーディネーションを基にした被服コーディネーションマップ

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

Yaguchi, Yuichi, and Ryuichi Oka. "Spherical visualization of image data with clustering." JACIII, vol. 17, no. 4, pp. 573-580, 2013.

〔学会発表〕(計 11 件)

Banage T. G. S. Kumara, Incheon Paik, Hiroki Ohashi, Yuichi Yaguchi, Wuhui Chen, Context Aware Filtering and Visualization of Web Service Clusters, Proceedings of International Conference on Web Service

(ICWS 2014), Alaska, U.S.A. June 2014.

Kumara, B.T.G.S.; Paik, I.; Ohashi, H.; Yaguchi, Y., "Web service filtering and visualization with context aware similarity to bootstrap clustering," Awareness Science and Technology and Ubi-Media Computing (iCAST-UMEDIA), 2013 International Joint Conference on ,pp.220,226, 2-4 Nov. 2013 doi: 10.1109/ICAwST.2013.6765437

Kumara, Banage TGS, et al. "Clustering and Spherical Visualization of Web Services." Services Computing (SCC), 2013 IEEE International Conference on. IEEE, 2013.

Wada, S.; Yaguchi, Y.; Ogata, R.; Wadanobe, Y.; Naruse, K.; Oka, R., "Associated Keyword analysis for temporal data with spatial visualization," Awareness Science and Technology and Ubi-Media Computing (iCAST-UMEDIA), 2013 International Joint Conference on , vol., no., pp.243,249, 2-4 Nov. 2013 doi: 10.1109/ICAwST.2013.6765441

Sasaki, Takeshi; Yaguchi, Yuichi; Watanobe, Yutaka; Oka, Ryuichi, "Extracting a spatial ontology from a large Flickr tag dataset," Awareness Science and Technology (iCAST), 2012 4th International Conference on , vol., no., pp.91,97, 21-24 Aug. 2012 doi: 10.1109/ICAwST.2012.6469595

Yaguchi, Yuichi; Oka, Ryuichi, "Spherical visualization of image data with clustering," Awareness Science and Technology (iCAST), 2012 4th International Conference on , vol., no., pp.200,206, 21-24 Aug. 2012 doi: 10.1109/ICAwST.2012.6469614

日野 陽平, 矢口 勇一, 轟 智則, “服の形状や色に基づく類似度を用いた球面マップ生成”, 可視化情報学会全国講演会 2013 会津, B109, 2013.

矢口 勇一, 轟 智則, 古沢 宏太, 日野 陽平, “Style Share: 球面上に展開する被服コーディネーションマップの生成”, 可視化情報学会全国講演会 2013 会津, B111, 2013

廣戸 裕大, 矢口 勇一, 渡部 有隆, 岡 隆一, “ソースコード類似性に基づく Aizu Online Judge の類似アルゴリズムマップの生成”, 可視化情報学会全国講演会 2013 会

津, B113, 2013,

和田 俊輔, 矢口 勇一, 尾形 亮, 渡部 有隆, 成瀬 継太郎, 岡 隆一, “空間可視化を用いた時系列データの連想単語解析”, 可視化情報学会全国講演会 2013 会津, D109, 2013.

矢口 勇一, 鷲山 英喜, “津波による遺失物写真の返却に資する球面上への画像クラスタリング”, 可視化情報学会全国講演会 2013 会津, E201, 2013. (イマジン)

E

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

<http://iplab.u-aizu.ac.jp>

6. 研究組織

(1) 研究代表者

矢口勇一 (会津大学)

研究者番号：00609109.