

## 「何時、何処で、誰と」3つのメタ情報を用いた大量個人画像の一覧可視化手法

五味愛<sup>†</sup> 伊藤貴之<sup>†</sup>

本報告では、「何時、何処で、誰と」というメタ情報を基に、直感的に大量画像を閲覧・分析することができる個人画像を対象にした3次元画像ブラウザを提案する。本ブラウザは、撮影場所と時間を基に自動でクラスタリングした画像を3次元上に閲覧することができるウィンドウと、イベントの共有者によってクラスタリングされた被写体情報を表示するウィンドウを、互いに連携操作できる形で表示する。この仕組みによって本ブラウザは、大量の個人画像から、画像所有者の生活ログの分析ツールとしても機能する。

### “When, Where, Who” Visualization Of Personal Photo Data By Using 3 Metadata

AI GOMI<sup>†</sup> TAKAYUKI ITOH<sup>††</sup>

This paper presents a personal photograph browser, which intuitively displays photographs from their shooting locations and times, as well as persons, and assists the analysis of history of photograph owners. It supports linked two views; one of the views displays photographs clustered based on their locations and times in 3D space, and the other displays persons clustered based on their co-occurrence in the events. This browser can be a life log analysis tool to explore the events of photograph owners.

## 1. 概要

近年のデジタルカメラの急速な発達に伴い、個人が所有するデジタル画像は、何千何万枚と増加の一途を辿っている。個人画像は人々の思い出を含んでいることが多いが、我々の画像に対する記憶は日が経つにつれ薄れていく。それゆえ、思い出の整理という意味でも画像の整理・検索は重要であると考えられる。また、画像の閲覧技術は、大量画像の整理や検索において重要な課題となってきた。個人画像において、重要な画像情報とは、「いつ、何処で、誰と」という点であると筆者は考える。例えば、ある画像を見た時、連想することは、「春頃の長野で行われた姉の結婚式」というように、曖昧ではあるが、「いつ、何処で、誰と」という情報は、個人の記憶に残りやすく、各々の画像から連想しやすい情報であると言える。

本報告では、撮影日時、撮影場所、被写体情報を基に、大量画像を「○○旅行」や「××パーティ」といったイベントに自動分類し、その結果であるイベントごとに画像を一覧表示する3次元画像ブラウザ MIAOW(Memorized Images As Organized by When/Where/Who)を提案する。MIAOWは、LTViewとPViewの2つの連携されたウィンドウを持つ(図1)。LTViewは、x軸とy軸に経度と緯度、z軸に撮影時間を割り当てられている。画面空間充填による階層型データの可視化手法を取り入れた画像ブラウザ CAT(Clustered Album Thumbnail)[3]のアルゴリズムを拡張し、3次元上にイベントごと画像配置を行っていく。PViewでは、イベントごとに被写体の共起性を調べ、被写体のグループを作成し、そのグループごとに被写体の顔画像を表示する。またMIAOWでは、ユーザのマウス操作によって簡単に、回転・移動・拡大縮小操作を行うことができる。そして画像を選択することで、その画像の撮影場所、撮影日時、被写体情報を全体の情報と比較しながら閲覧することができる。これにより、画像の直感的な検索だけでなく、個人の間関係や行動範囲などの分析も行うことができる。例えば、「○○さんとは3年以上会っていないから、○○さんを含めたイベントを企画しよう。」といった新たなイベントの企画支援に繋がる分析が行えると考える。本報告では、個人画像の分析におけるMIAOWの有用性を検証する。

## 2. 関連研究

### 2.1 画像ブラウザ

既に多くの個人画像ブラウザが商用ソフトウェアとして開発されている。個人用の

図1 MIAOW の概要

画像ブラウザとしては、Picasa[1](google)やiPhoto[2] (Apple)などが有名である。これらは、撮影場所、撮影日時、人物などの情報から画像を整理閲覧することができる。また、ネットサーバと連携させ、ウェブ上の画像共有コミュニティサイトへ簡単に画像をアップロードすることができる。クエリ機能も充実している。

大量画像ブラウザに関する研究の多くは、画像間の類似性を考慮した画像サムネイルのレイアウトに着目している。その中でも初期の多くの手法では、多次元尺度構成法(MDS) や主成分分析(PCA) を活用し、低次元ユークリッド空間に画像サムネイルを配置している。また、木構造やグラフ構造などを対象とした情報可視化技術のいくつかは、大量画像ブラウザの研究にも適用されている。Jankun-Kellyらは、局所拡大表示型のグラフレイアウト技術を大量画像ブラウザに適用している[7]。Bederson らは、2.2節で紹介するQuantum Treemap[6]という階層型データ可視化手法を採用した大量画像ブラウザPhotoMesa[4]を提案している。画像注釈に基づく大量画像の閲覧手法として、YangらはSIB(Semantic Image Browser) [8]を提案している。五味らは、階層的に画像をクラスタリングし、生成された各クラスタを入れ子型の長方形で表し、画面空間充填の配置を行うアルゴリズムを適用した大量画像ブラウザCAT[3]を提案している。本報告では、このCATのアルゴリズムを取り入れ、画像クラスタを入れ子型の長方形で表

し、メタ情報を基に長方形の配置を行う。

## 2.2 階層型データの可視化

階層型データの情報可視化手法の中でも有名なものに、木を表示する手法と、画面空間充填による手法があげられる。MIAOWの先行研究であるCATは、画面空間充填手法に属する。CATは画面空間充填による階層型データ可視化手法「平安京ビュー」[5]をベースにした手法である。「平安京ビュー」は、入れ子型の長方形の枠で階層を表し、葉ノードをアイコンで表示する。「平安京ビュー」は階層構造の最下位に属する葉ノード群を、クリックしやすく等しいサイズのアイコンとして、しかも互いに重ならずに表示できる、という特徴がある。このアイコンを画像サムネイルに置き換えることで、CAT は大量の画像を、等しいサイズで互いに重ならずに一覧表示できる。2.1節で紹介したPhotoMesa[4]が用いているQuantumTreemap[6]も、画面空間充填による階層型データ可視化手法であり、「平安京ビュー」に類似した特徴を持つ。文献[5]ではQuantum Treemapと「平安京ビュー」の比較評価結果を掲載している。この比較評価結果において「平安京ビュー」は、類似する階層型データのレイアウトの安定性と、長方形領域の縦横比において、よい結果を得ている。またQuantum Treemapと違って「平安京ビュー」は、不均一な2段階以上の階層を有する任意の階層構造を表現できる。これらの長所は、そのままCATのPhotoMesaに対する優位性としてあげられる。なお、文献[5]で述べるように「平安京ビュー」は、画像配置の理想位置を記述したテンプレートによって、長方形の位置を操作する機能を有する。MIAOWでは、このテンプレート機能を活用し、撮影位置情報を基に、地図上の位置を反映させ配置を行う。

## 3. 提案内容

### 3.1 大量画像クラスタリングの処理手順

本報告で扱う画像は、撮影日時、撮影場所そして被写体情報をメタ情報として含む画像を想定している。撮影日時、撮影場所情報は予め画像に付加されているメタデータから取得する。被写体情報は、顔認識技術を使用して顔の識別を行い、各画像から取得する。画像 $P_i$ の属性を以下のように定義をする。

$$P_i = (\lambda_i, \varphi_i, t_i, f_{1i}, f_{2i}, \dots, f_{m_i}) \quad (1)$$

ここで、 $\lambda_i$ は $P_i$ の経度、 $\varphi_i$ は $P_i$ の緯度、 $t_i$ は $P_i$ の撮影日時、 $f_{ji}$ は $P_i$ の被写体を示し、 $m_i$ は $P_i$ の被写体数を示す。

本実装では、前処理として画像と被写体のクラスタを生成する。以下にクラスタリングの処理手順を述べる(図2)。

**[手順1-1]**位置情報を基に、すべての画像に対しクラスタリングを行い、高階層クラスタを生成する。

**[手順1-2]**撮影日時情報を基に、各高階層クラスタの画像に対し、クラスタリングを行

い、低階層クラスタを生成する。ここで生成されたクラスタをイベントと定義する。

**[手順1-3]**各イベントにおける被写体の共起性から、被写体のクラスタリングを行う。

本手法では位置情報によってXY平面上の全体的な配置を決定するため、最初に位置情報に基づいて画像のクラスタリングを行う。これにより本手法では、XY平面上に、地理的位置に近いイベントをまとめて表示することができる。

手順1-1では、2枚の画像間の距離を撮影場所における物理的距離と定義し、経度、緯度情報を基にデンドログラムを生成する。そして、任意の閾値により、画像を高階層クラスタに分類する。ここで、生成された高階層クラスタのメルカトル地図上での理想値を求める。我々の実装では、クラスタに属する画像のメルカトル地図上の位置の平均値を理想値として定義する。これら理想値は、3.2節で述べるクラスタ配置の際に使用する。

手順1-2では、撮影日時を基にクラスタリングを行う。我々の実装では、各高階層クラスタの画像を撮影日時順に並び替え、2枚の画像間の距離を撮影時間間隔として、デンドログラムを生成する。そして、任意の閾値によって低階層クラスタを生成する。このとき、各クラスタの被写体情報を集計し、距離の近いクラスタにおいて、被写体情報の共起性が高ければ結合を行う。クラスタリングを行った後、各低階層クラスタの理想値を計算する。我々の実装では、画像の撮影日時における平均値を理想値と定義する。これら理想値は、3.3節で述べるクラスタ配置の際に使用する。

手順1-3では、誰がどのイベントに参加したかという被写体情報から、被写体のクラスタリングを行う。 $i$ 番目のイベントに参加した被写体 $j$ の画像枚数を $q_{ij}$ とすると、被写体 $j$ は $n$ 次元ベクトル  $(q_{1j}, q_{2j}, \dots, q_{nj})$  で表すことができる。ここで、 $n$ はイベント数を表す。本手法では、任意の被写体間のユークリッド距離を $n$ 次元ベクトルから計算し、この距離に基づいて被写体のデンドログラムを生成する。そして任意の閾値によって被写体のクラスタを生成する。結果的にこのクラスタは、家族や友人や同僚といったグループを表す。

### 3.2 位置情報を基にしたXY平面への画像配置

本手法では以下の処理手順により、XY平面に大量画像を画面配置する。なお本手法では、2.2節で述べたとおり、空間充填による階層型データ可視化手法「平安京ビュー」と同じ画面配置アルゴリズムを導入している。

**[手順2-1]**低階層クラスタに属する画像サムネイルを格子状に配置し、長方形でそれらを囲む。すべての低階層クラスタに、同様の過程を繰り返す。

**[手順2-2]**低階層クラスタを表現するすべての長方形を隙間なく配置し、高階層クラスタを表現する長方形で囲む。すべての高階層クラスタに同様の過程を繰り返す。

**[手順2-3]**高階層クラスタを示す長方形を、撮影場所を考慮しながら配置を行う。MIAOWは「平安京ビュー」と同様、文献[5]に示す以下の値を最小化するように長方形群を配置する。

$$aA + rR + dD \quad (3)$$

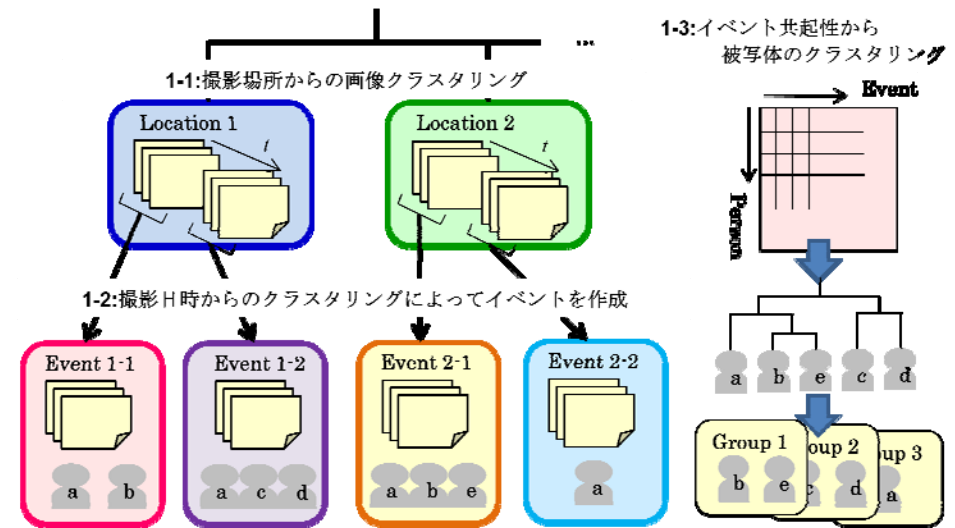


図2 画像と被写体のクラスタリング

ここで $a, r, d$ は、正の定数でありユーザが指定する。 $A$ は長方形配置における前後間の長方形の面積比、 $R$ は長方形配置における前後間の長方形の縦横比、 $D$ は長方形の実際の位置とテンプレートに描かれた理想的な位置との間の距離である。本手法では、変数 $R$ を、理想的な縦横比(=代表画像の縦横比)に対する長方形領域の縦横比の悪化の度合い、として計算する。理想的な縦横比とは、手順2-1においては、代表画像の縦横比とし、手順2-2においては、正方形の縦横比すなわち1とする。手順2-3においては、メルカトル地図の縦横比とする。また、変数 $D$ を、実際の長方形の位置と手順1-1で求めた地図上の理想位置との間の距離とする。手順2-1、2-2においては $d=0$ とし、手順2-3では、 $d$ に任意の値を指定する。すなわち、高階層クラスタのみ、地図上の理想位置に近くなるように配置をする。

### 3.3 日時情報を基にしたXZ平面への画像配置

XY平面に入れ子型に長方形を配置した後、低階層を示す長方形に撮影日時を基に $z$ 値を与え、XZ平面への配置を行う。我々の実装では、撮影場所を示すXY平面と撮影日時情報を示すXZ平面の切り替えを、ユーザのマウス操作によってスムーズに行うことができる。XY平面からXZ平面に移動する際、XY平面上において求めた $x$ 座標値を固定しながら、XZ平面に切り替える。XZ平面上の配置アルゴリズムを図3に示す。

**[手順3-1]**手順2-2において計算された低階層クラスタを示す長方形のx座標値を固定したまま、低階層クラスタを撮影日時順にXZ平面上に配置する。

**[手順3-2]**低階層クラスタを示す各長方形の重複を避けながら、z座標値を修正する。

**[手順3-2]**各画像サムネイルのz座標値を、低階層クラスタを示す長方形内に収まるように修正する。

手順 3-1 において、すべての低階層クラスタの平均撮影時間からクラスタを昇順にソートし、XY 平面における x 座標値を保持しながら、重複を避けながら長方形を XZ 平面に配置する。ここで z 座標値の最小値を  $z_{min}$  と定義し、以下の条件を満たすように、長方形  $R$  の配置を行う(図 3 左)。

1. すでに配置されている長方形に重複しないように  $R$  を配置する。
2. すでに配置されている長方形に隣接するか、 $z=Z_{min}$  として配置する。

手順 3-2 では、長方形の z 値を修正する(図 3 右)。手順 3-1 ですべての長方形を配置した際に、最大となる Z 値を  $Z_{max}$  として各長方形の理想値  $Z_i$  を以下の式で求める。

$$Z_i = \frac{t_i - t_{min}}{t_{max} - t_{min}} (Z_{max} - Z_{min}) + Z_{min} \quad (4)$$

ここで  $R_i$  の撮影時間を  $t_i$  とし、 $t_{max}$  と  $t_{min}$  を低階層クラスタにおける撮影時間の最大値と最小値とする。長方形の重複を避け、理想値  $Z_i$  に近づくように z 値を修正する。

図 3 XZ 平面における長方形の配置手順

### 3.4 LTViewとPViewの連携

LTViewは、マウスの左ボタンに移動、右ボタンに回転、ホイールにズーム操作を割り当てている。回転操作により、x軸を中心として回転させることで、容易にXY平面とXZ平面を切り替えられる。その際、画像は視線方向に平行な法線方向を保って表示されるので、回転中でも画像を常に目で追うことができる。また、ズーム操作では、

ズームアウト時は、各低階層クラスタの代表画像を表示させ、ズームイン時は、個々の画像サムネイルを表示する詳細度制御を設けている。画面上に表示する画像のみをメモリに読み込むことによって、表示速度の高速化を実現している。これにより何万枚といった画像セットを扱うことが可能である。

PViewは、「平安京ビュー」の画面配置アルゴリズムにより、処理1-3で求めた被写体のクラスタごとに顔画像を表示する。多くの場合において、画像所有者にとって意味あるグループ（家族、友人、同僚など）を単位とした被写体表示を実現できる。

LTViewとPViewは連携されており、LTView上のイベントの代表画像とタイムライン上、そして、PView上の顔画像を選択することができる(図4)。LTView上で、ユーザがイベントの代表画像をクリックすると、タイムライン上がハイライトされるとともに、PView上においてそのイベントを共有している被写体がハイライトされる。また、PView上で、ある被写体を選択すると、その被写体が参加したイベントとタイムライン上の特定の部分がハイライトされる。このように、連携されたLTViewとPViewを操作することで、ユーザは簡単に画像の分析を行うことができる。

図4 LTViewとPViewの連携操作

## 4. 実行例

我々はWindowsXPおよびJava JDK 1.6.0を用いて本手法を実装し、IBM ThinkPad T500 (CPU 2.8GHz, RAM 2GB)上にて本手法を実行した。使用した画像セットは、個人が所有しているJPEG画像4804枚から、サイズ100×75画素のサムネイルと、サイズ任意の代表画像を生成したものである。上記の画像セットにおいて、前処理のクラスタリングでは、高階層クラスタが38個、低階層クラスタ(イベント数)が73個生成された。

### 4.1 実画像セットに基づくケーススタディ

本実験で扱う画像セットは、ある女子大学生が2004年から2009年に実生活の中で撮影した個人画像に基づくものである。図5は、2004年から2009年の各年を選択した結果を示している。2004年において彼女は主に、ヨーロッパや太平洋の島国で大学友人や家族とイベントを共有している(図5左上)。2005年には、研究室メンバーや家族と共に多くのイベントを共有している。実際に2005年は、彼女にとって研究室配属1年目であり、研究室メンバーとの交流が始まった年でもある(図5中央上)。2006年には、大学友人との交流が再び行われ、家族との交流も頻繁に行われている(図5右上)。そして、2007年には、アメリカのさまざまな場所に訪れていることがわかる(図5左下)。2007年を他の年と比べてみると、2007年以外はヨーロッパに行き、この年だけはアメリカに行っていることも確認できる。そして2008年には、多くの同僚とイベントを共有していることがわかる(図5中央下)。実際に2008年には、画像所有者は社会人になり、さまざまな出会いがあった。2009年には、海外での写真がほとんどない(図5右下)。この年は実際に忙しく、海外に行く時間がなかったこともうかがえる。このように本手法では、年ごとのイベントを場所と人物を同時に表示することで、画像所有者の生活ログを容易に分析することができる。

図6は、同じグループに属する人物を分析した結果を示す。この結果から、画像所有者と友人Eは、2006年以降イベントを共有していないことがわかる(図6左)。一方で、グループ内の他の友人達の結果は、図6右の友人Bの結果とほぼ同じであり、2008-2009年にも3回イベントを共有している。画像所有者はこの結果をみて実際に、友人Fもこれらのイベントに招待してもよかったのかもしれない、友人Fを招待し新たなイベントの計画を立てるべきだ、と考えた。本手法ではこのように、人間関係を分析することで、長年交流がない人物を容易に発見することができる。

図7は、画像所有者の姉との関係を可視化した結果である。姉と共有した最後のイベントは2007年であった。画像所有者は、2007年以降なぜ姉とイベントを開催していないのか疑問に思った。そこで最後のイベントをズームしてみると、姉の結婚式であることがわかった。画像所有者は、姉家族と家族ぐるみのイベントの計画を立てようと思った。このようにズーム操作によって、イベントの詳細な画像を簡単に閲覧することができ、イベント分析が容易になる。

図5 タイムラインから各年を選択した結果

(左上2004年, 中央上2005年, 右上2006年, 左下2007年, 中央下2008年, 右下2009年)

図6 大学友人グループのイベント比較(左友人Aを選択, 右友人Bを選択)

図7 特定人物を選択した結果

#### 4.2 ユーザ実験

被験者に前述の画像セットを使ってMIAOWを操作してもらい、ブラウザから読み取れる内容を分析してもらった。被験者は情報科学科の学生10人(画像所有者ではない)であった。以下がユーザ実験によって得られた回答内容である。

- 画像所有者は 2007 年にのみアメリカに訪れている。一方で、それ以外の年は、ヨーロッパに毎年訪れている。
- 画像所有者は 2008 年以降同僚と多くのイベントを共有している。
- 画像所有者は特定の人物と海外に訪れている。他の人物は国内でのみイベントを共有している。
- 画像所有者は大学友人と頻繁にイベントを共有している。
- 画像所有者は毎年家族とイベントを共有している。
- 2008 年以降に幼児がイベントに参加し始めた。

これらの回答から、どの被験者も的確に素早く、画像所有者のイベントの特徴を理解していることがわかった。この結果から、本ブラウザを用いることで、個人の生活ログの分析が容易になることがわかった。ブラウザの操作が簡単であることから、興味深い場所、時間、人物から画像を閲覧することによって、直感的にイベントの特徴をとらえることができると考えられる。特に、人物がグループに分かれていることは有益な情報であるという意見も出た。

同時に、被験者に対しブラウザの改良点についても以下のような意見を貰った。

- 人物の登場回数によって色もしくは画像サイズを変更すれば、より情報を得ることができる。

- 世界地図上で選択が可能であれば、より使いやすい。
- 色によって国や地域が分けられると更にわかりやすい。
- 時間選択が四季や月ごとに分かっていると更によい。
- 選択されたイベントのみ表示をしたほうが見やすい。

これらは、我々の実装において容易に拡張できると想定されるので、今後の課題として取り入れたい。

#### 5. まとめ・今後の課題

本報告では、場所、時間、被写体情報を基に個人画像を閲覧分析することができるブラウザを提案した。本手法では、場所と時間情報から画像セットをクラスタリングし、イベントを生成する。そして、イベントの共有者を元に被写体のクラスタリングを行う。このクラスタリング結果を表示するブラウザは、場所と時間を基に3次元空間上に画像を配置する LTView 画面と、グループごとに被写体を配置する PView 画面を持ち合わせている。ユーザが画像を探索する際には、この2つの画面を連携操作することができる。本報告では、実際の個人写真におけるケーススタディを報告し、ユーザ実験を行った。

我々の行った前処理の実装は、代表画像の選択以外はほぼ自動で行うことができる。代表画像の選択方法において、何人かの被験者から、撮影時間や、閲覧回数の多い画像や、被写体数の多い画像などを代表画像に選択したいとコメントをもらった。これらに関しては、今後実験を行いたい。また、何万枚といった膨大な画像を使用した実験とユーザ評価を行いたい。

#### 参考文献

- 1) Google Picasa. 2009. Information at <http://www.picasa.com/>.
- 2) Apple iPhoto. 2009. Information at <http://www.apple.com/iphoto/>.
- 3) 五味, 宮崎, 伊藤, Li, CAT:大量画像の一覧可視化と詳細度制御のための GUI, 画像電子学会誌, Vol. 38, No 4, pp. 436-443, 2008.
- 4) Bederson B., B., PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps, User Interface Software and Technology, pp. 71-80, 2001.
- 5) 伊藤, 山口, 小山田, 長方形の入れ子構造による階層型データ視覚化手法の計算時間および画面占有面積の改善, 可視化情報学会論文集, Vol. 26, No. 6, pp. 51-61, 2006.
- 6) Bederson B., Schneiderman B., Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies, ACM Trans. on Graphics, Vol. 21, No. 4, pp. 833-854, 2002.
- 7) Jankun-Kelly T. J., Ma K.-L., MoireGraphs: Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes, IEEE Information Visualization, pp. 59-66, 2003.
- 8) Yang J., et al., Semantic Image Browser: Bridging Information Visualization with Automated Intelligent Image Analysis, IEEE Visual Analytics in Science and Technology, 2006.

## 著者紹介



### 五味愛

2006年お茶の水女子大学理学部情報科学科卒業。2008年お茶の水女子大学大学院人間文化研究科博士前期課程修了。現在お茶の水女子大学大学院人間文化研究科博士後期課程在学中。



### 伊藤貴之（正会員）

1990年早稲田大学工学部電子通信学科卒業。1992年早稲田大学大学院理工学研究科修士課程修了。同年日本アイ・ビー・エム(株)入社。1997年博士(工学)。2000年米国カーネギーメロン大学客員研究員。2003年京都大学大学院情報学研究科COE研究員(客員助教授相当)兼任。2005年お茶の水女子大学理学部情報科学科准教授。ACM, IEEE CS, 他会員。