

空間関係を利用した写真と注釈の相補的な共有モデル

藤田 秀之 有川 正俊

東京大学空間情報科学研究センター 〒153-8904 東京都目黒区駒場 4-6-1

E-mail: {fujita, arikawa}@csis.u-tokyo.ac.jp

あらまし 写真集合の検索手法として、写真に写る内容をテキストで記述したラベルをキーとするものが一般的であるが、すべての写真にラベルを入力するのはユーザにとって負荷が高く、また写真は入力済みのラベルでしか検索されないといった問題点がある。一方で GPS 等の空間センサの普及で、撮影時の位置や方向等をメタデータとして持つ写真が流通することが想定される。本論文では、このような写真を対象としたラベル共有システムを提案する。本システムはまず、複数の写真に入力済みのラベルを収集し、各ラベルが指し示す対象の地理的位置を決定する。続いて、各ラベルの対象位置と、写真集合内の各写真の視点および撮影方向との空間的な関係から、写真にラベルの対象が写っていると考えられる場合、そのラベルを伝播させる。プロトタイプシステムでは写真集合内でのラベルの共有、および写真上の適切な位置へのラベルの配置を実現した。

キーワード 時空間 DB, ユーザインタフェース, GIS, ユビキタスマッピング

Complementary Sharing of Photographs and Annotations Considering their Spatial Relationships

Hideyuki FUJITA Masatoshi ARIKAWA

Center for Spatial Information Science, University of Tokyo 4-6-1, Komaba, Meguro-Ku, Tokyo 153-8904, Japan

E-mail: {fujita, arikawa}@csis.u-tokyo.ac.jp

Abstract Digital photographs are usually organized by text labeling in order to search or browse a collection. But labeling many photographs costs users much effort. On the other hand, many kinds of spatial sensors including GPS have become popular today. We assume there will be many photographs with spatial metadata such as viewpoints and view directions. This paper suggests label sharing system for such photographs. The system (i) stores label data submitted by users, and (ii) decide locations of objects pointed by stored label texts. When a photograph requires label data, the system (iii) selects and shares all label data which are judged to be seen in the photograph by considering field of view (FOV) of the photograph. In this framework, label texts are placed on appropriate positions on photographs, and that makes photographs clickable when label data have URLs.

Keyword Spatio - Temporal Database, User Interface, GIS, Ubiquitous Mapping

1. はじめに

携帯電話に搭載される例を含めてデジタルカメラが普及しており、多くのユーザが日常的に写真を撮影し、一部は WEB 上に流通している。また、同じく携帯電話に搭載され人ナビに利用される例にみられるように、GPS や電子コンパス等の空間センサが個人向けのデバイスとして普及し始めている。これらの空間センサを利用することで写真を撮影しながら同時にカメラの位置や方向を取得できる。また、デジタルカメラで撮影した写真ファイルのメタデータの業界標準として、日本電子工業振興協会が定める Exif format[1]が日本国内外を問わず普及しており、国内の民生用デジタルカメラのほぼ 100%が対応している。対象となるメタデータには、基本的な撮影条件（撮影日時、シャッ

タースピード、絞り、ホワイトバランス等）に加え、空間的なパラメータ（撮影位置、撮影高度、撮影方向、測地系、GPS 精度等）も含まれる。Exif 情報は JPEG や TIFF といった一般的な写真ファイルのヘッダー部分に記録されるため、写真ファイルが空間メタデータと一体となって流通することになる。GPS 等の空間センサを搭載し、写真ファイルに空間的なメタデータを EXIF format で記録するデジタルカメラも製品化されている。以上より、今後、位置や方向等の空間的なメタデータを埋め込まれたデジタル写真データが WEB 上で流通することが想定される。筆者らはこのような写真を空間スキーマ付き写真と名付け、空間的なメタデータをキーとした空間的な検索手法や、他の空間データとの重ね合わせの手法を提案してきた[2][3]。一方

で、例えば「東京タワーが写っている写真が欲しい」というような、写真に写る内容によってなされる検索を処理する手法として、写真に写る内容を、例えば「東京タワー」というようにテキストで表現したラベルデータを各写真データに関連付け、このラベルデータをキーとして写真を扱うというものが一般的である。複数の画像検索に関する研究[4]や、WEB上の画像検索サイト[5]、市販のアルバムソフト[6]がこの枠組みを採用しているが、既存手法には、次の問題がある。

- ・ ラベルテキストを持たない写真は検索できない
- ・ 大量の写真に対してラベルテキストを手入力するのは負荷が高い

他方で、近年、飲食店やランドマーク等の位置情報がPOI (Point of Interest) データとして整備されており、カーナビゲーションやヒューマンナビゲーション、個人向けの地図サービス等で利用するコンテンツとして重視されている。POI データは、一般に店舗名等のテキストとその地理的位置座標（東経、北緯等）を主要素とする。例えば、位置座標を介してラベルとして地図上の適切な位置に重ねて表示したり、テキストや音声によるナビゲーションにおいて道案内文章を自動生成する際の語要素にするといった利用がなされている。地図データをはじめとする空間データおよび空間コンテンツの記述・通信・交換に関するプロトコルをWEB上での流通を目的として定めたJIS規格であるG-XMLでは、POI データに店舗の電話番号や営業時間等さまざまな要素が追加されることを想定して、POI データ作成者が要素を任意に定義でき、さらにその交換可能性も保障するための辞書付きPOIの概念も導入されており、POI データ共有の枠組みが整いつつある[7]。

以上を背景とし、本論文では、不特定多数のユーザー間で空間スキーマ付き写真に対するラベルテキストを共有するラベル共有システムを提案し、前述のテキストラベルをキーとした写真検索における問題点の解決を試みた。本論文は以下を前提としている。

- ・ 対象とする写真は空間スキーマ付き写真である
- ・ ユーザは写真上で対象の写る位置を指定してラベルテキストを入力する。つまりラベルデータは特定の写真上でのピクセル座標を持つ
- ・ 1枚の写真が複数のラベルを持つ
- ・ POI データは地理座標を持つ

ラベル共有システムの動作は次のとおりである。

- (1) 空間スキーマ付き写真に対して入力されたラベルデータをラベル共有システムに蓄積する
- (2) 共有データベースのラベルデータをPOIデータ化する
- (3) ラベル共有システムにラベルデータを要求した空間スキーマ付き写真に対しては、共有データベー

スのPOIデータのうち、指し示す対象がその写真に写るものすべてをラベルデータ化し伝播させる。通常、特定の写真のメタデータとして扱われるラベルデータを、写真データとは独立した、地理的位置座標を持つPOIデータに変換することで、その再利用性を向上させる。例えばラベルデータを持たない写真データに対してもラベルデータが伝播する。動作(2)は、ラベルが指し示す対象の地理的位置を決定することで実現されるが、次の問題がある。

- ・ 空間スキーマ付き写真に対してラベルデータが記述された場合でも、そのラベルの位置として写真がメタデータとして持つ、GPS等で取得したカメラの位置（視点）が適切とは限らない。

理由として、空間スキーマ付き写真に対して記述されたラベルデータが指し示すのは、その写真の撮影対象であるが、GPS等によって取得される、空間スキーマ付き写真がメタデータとして持つ位置は、撮影時のカメラの位置（視点）であり、撮影対象の位置ではないことが挙げられる。そこでラベルが指し示す対象の位置を決定する手法を提案する。また(3)に関して、前述の既存POIデータも写真のラベルデータに変換できる。さらに、写真データへ伝播した複数のラベルテキストを、写真上の適切な位置へ重ねあわせて表示した。これはPOIデータを地図上だけでなく、写真上に配置する枠組みである。提案手法に関して、次のような利用環境が想定される。「GPSと電子コンパスを搭載したデジタルカメラで、渋谷駅八チ公前交差点付近の空間スキーマ付き写真を撮影した。ラベル共有システムにアクセスし、『Q-FRONT』『センター街入口』というラベルデータを取得した。写真上に表示させると、白い建物とその脇のゲート上にそれぞれ『Q-FRONT』『センター街入口』というテキストが配置された。建物上に配置された『Q-FRONT』というテキストをクリックすると商業ビルである『Q-FRONT』の公式サイトが表示された。次に、写真上をクリックし『西村フルーツパーラー』というテキストラベルを入力し、ラベル共有システムに登録した。以前他のユーザにより同一のラベルが登録されていたため、このラベルはPOIデータされ、他の空間スキーマ付き写真にも伝播することになる。」

本論文の構成は次の通りである。まず、対象データである空間スキーマ付き写真、ラベルデータおよびPOIのデータ構造と、提案するラベル共有システムの構成をまとめる(2章)。これらを前提として、同一対象を指し示す複数のラベルデータから対象の地理的位置を決定する手法(3章)と、各空間スキーマ付き写真に対して、その写真に写っていると推定されるラベルデータを伝播させる手法(4章)を説明する。次に、

提案手法を実装したプロトタイプシステムを紹介し、最後に結論と今後の課題をまとめる(5章)。

2. データおよびシステム構成

本章では、空間スキーマ付き写真とラベルデータ、POI データに関して、そのデータ構造と幾何学モデルを定義し、次にラベル共有システムの構成と、処理手順およびデータの流れを説明する。まず、本論文が扱う空間スキーマ付き写真データ P は次の構造を持つ。

$$P = (I, SP, L) \quad : \text{空間スキーマ付き写真データ}$$

ここで、写真イメージ I は画像自体、 SP は写真が持つ空間スキーマを示し、ラベル集合 L は写真に対応付けられたラベルデータの集合である。空間スキーマ SP は以下のように構成される。

$$SP = (S, V, \theta_w, \theta_h) \quad : \text{空間スキーマ}$$

ここで視点 $S = (X_S, Y_S, Z_S)$ は写真撮影時のカメラの位置であり、撮影方向 $V = (X_V, Y_V, Z_V)$ は写真撮影時のカメラの視線の方向ベクトルである。それぞれ絶対座標(地理座標)で表現する。画角 θ_w, θ_h は写真が切り取る空間の縦横の幅を決定する。異なるカメラ間でのレンズの焦点距離の違いや同一のカメラにおけるズーム値の変化が、画角の変化に対応する。各パラメータの空間的な関係を図1に示す。

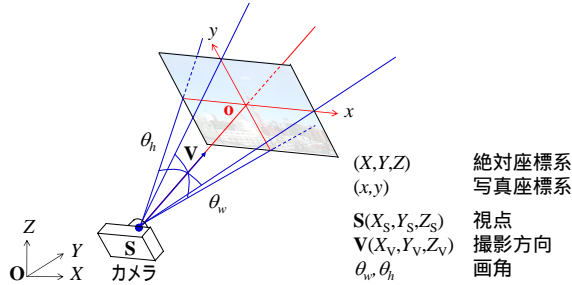


図1. 空間スキーマ付き写真

続いて、ラベル集合 L はラベルデータ l の集合である。

$$L = \{l_1, l_2, l_3, \dots\} \quad : \text{ラベル集合}$$

ラベルデータ l は次の構造を持つ。

$$l = (text, m, url) \quad : \text{ラベルデータ}$$

ラベルテキスト $text$ はラベルが指し示す対象を表す文字列である。写真上位置 $m = (x_m, y_m)$ はラベルが指し示す対象の写真上での位置であり、写真座標(ピクセル)である。 url は関連する WEB サイトの URL であり、 $null$ 値でもよい。各パラメータの関係を図2に示す。本論文では、ユーザが写真上をクリックして写真上位置 m を指定し、ラベルテキスト $text$ を入力してラベルデータ l を作成することを想定している。また、1枚の写真が複数のラベルデータを持つこととする。

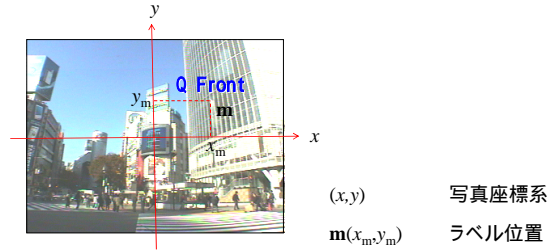


図2. ラベルデータ

続いて、POI データは次の構造を持つものとする。

$$POI = (text, M, url) \quad : \text{POI データ(地物データ)}$$

POI テキスト $text$ は POI が指し示す対象を表す文字列である。位置 $M = (X_M, Y_M, Z_M)$ は POI が指し示す対象の実世界での位置である。 url は関連する WEB サイトの URL であり、 $null$ 値でもよい。

次に、ラベル共有システムの構成は図3のとおりである。本システムの動作を以降にまとめる。

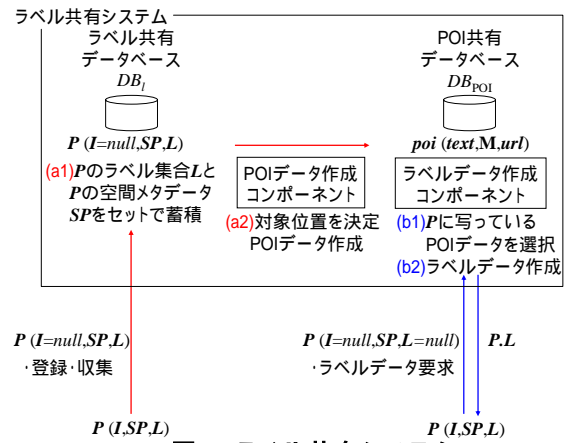


図3. ラベル共有システム

- ・新たなラベルデータの登録があった際：
 - (a1) 登録された写真データ P をラベル共有データベース DB_L に蓄積する。
 - (a2) DB_L に蓄積されたラベルデータの一部を POI データ化し、POI 共有データベース DB_{POI} に蓄積する。
 - ・ラベルデータの要求があった際：
 - (b1) ラベルデータを要求した空間スキーマ付き写真に対し DB_{POI} 中の POI データのうち、指し示す対象がその写真に写っているものすべてを選択する。
 - (b2) 選択された POI データをラベルデータ化しラベルデータ集合 L を作成して返す。
- 本システムにおいてクライアントとラベル共有システムの間で送受信されるデータは、ラベルデータを含む写真データ P であるが、写真イメージ I は送受信されない。つまり、(a1)のラベル共有データベース DB_L に受け渡され蓄積される P や、(b1)のラベル要求の際の引数としてシステムに渡す P に関して、 $P.I = null$ である。(a2)は POI データ作成コンポーネントが行う処理であ

り，詳しくは 3 章で説明する．(b1)(b2)はラベルデータ作成コンポーネントが行う処理であり，詳しくは 4 章で説明する．

3. 複数のラベルデータからの POI データ作成

前章で示したとおり，ラベルデータは写真上のピクセル座標を持つ，特定の写真に関するデータであるが，POI データは絶対座標（地理座標）を持つ，特定の写真とは独立したデータである．本章では，共有データベースに蓄積された複数のラベルデータを対象とし，図 4 に示すとおり複数の空間スキーマ付き写真が同一の対象を指し示すラベルを持つ場合，ラベルが指し示す対象の絶対位置を決定することで POI データを作成する．POI データの作成とは， $POI = (text, M, url)$ の各要素を決定することを指す．なお，本章では POI の指し示す対象の位置 $POI.M$ の算出に利用する各要素（空間スキーマ付き写真の空間スキーマ SP の要素である実世界座標系上の位置や方向等）が，高さ Z 座標を含む 3 次元座標を持つものとして議論するが， Z 座標を持たない 2 次元のデータを利用する場合は， $POI.M$ に関しても Z 座標を含まない 2 次元座標が算出される．

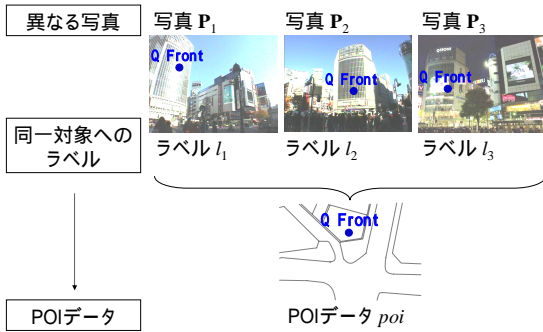


図 4. 複数のラベルデータからの POI データ作成

まず，同一対象を指し示す 2 つのラベルデータから POI データを作成する手順を示す．ラベル共有データベース DB_l に蓄積された，2 つの空間スキーマ付き写真 P_a ， P_b に関して以下のように定める．

$$P_a = (I_a, SP_a, L_a), \quad SP_a = (S_a, V_a, \theta_{wa}, \theta_{ha})$$

$$L_a = l_{a1}, l_{a2}, l_{a3}, \dots, l_{ap}, \dots$$

$$P_b = (I_b, SP_b, L_b), \quad SP_b = (S_b, V_b, \theta_{wb}, \theta_{hb})$$

$$L_b = l_{b1}, l_{b2}, l_{b3}, \dots, l_{bq}, \dots$$

複数のラベルデータが同一のラベルテキストを持つ場合，同一の対象を指し示すラベルデータであると判断する．すなわち， $l_{ap}.text = l_{bq}.text$ のとき，ラベル l_{ap} と l_{bq} の指し示す対象は同一である．このとき，

$$POI.text = l_{ap}.text = l_{bq}.text \quad (1)$$

と定める．以降 l_{ap} ， l_{bq} を l_a ， l_b と表す．続いて $POI.M$ を算出する．まず関数 $Direction$ を写真 P の空間スキーマ

マ SP ，写真 P に対するラベルデータ l を引数とし，写真 P の視点からラベル l への方向ベクトル $V_l = (X, Y, Z)$ を返すものと定義する．すなわち，

$$V_l = Direction(SP, l) \quad (2)$$

である．関数 $Direction$ による V_l の具体的な算出方法を説明する．一般に，図 5 に示すように絶対座標が $Q = (X, Y, Z)$ の点の，視点 S ，焦点距離 c ，撮影方向 $V = (X_v, Y_v, Z_v)$ の写真に上における写真座標が $q = (x, y)$ のとき，次の共線条件式(3)(4)が成立する[8]．

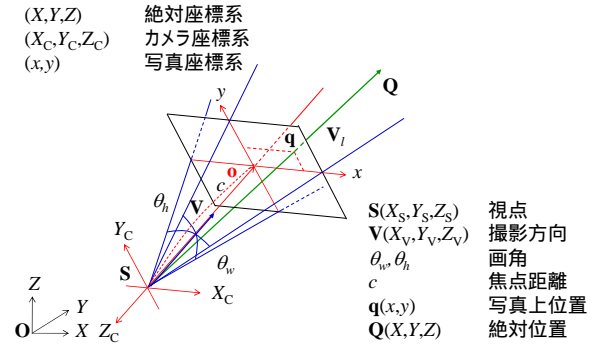


図 5. 共線条件式

$$X - X_s = (Z - Z_s) \frac{a_1 x + a_2 y - a_3 c}{a_7 x + a_8 y - a_9 c} \quad (3)$$

$$Y - Y_s = (Z - Z_s) \frac{a_4 x + a_5 y - a_6 c}{a_7 x + a_8 y - a_9 c} \quad (4)$$

ただし， R は撮影方向 V から定まる回転行列であり．焦点距離 c は式(6)より画角 θ_w と写真ピクセルサイズ横 w を用いて算出可能である．

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{pmatrix} = R \quad (5)$$

$$c = \frac{w}{2 \tan \frac{\theta_w}{2}} \quad (6)$$

式(3)(4)において，

$$\frac{a_1 x + a_2 y - a_3 c}{a_7 x + a_8 y - a_9 c} = f_1(x, y) \quad (7)$$

$$\frac{a_4 x + a_5 y - a_6 c}{a_7 x + a_8 y - a_9 c} = f_2(x, y) \quad (8)$$

$$Z - Z_s = f_3(x, y) \quad (9)$$

とおけば，式(3)(4)はそれぞれ次のように書ける．

$$X - X_s = f_1(x, y) f_3(x, y) \quad (10)$$

$$Y - Y_s = f_2(x, y) f_3(x, y) \quad (11)$$

視点 S から写真座標 m のラベル l が指し示す対象への方向 V_l は，式(9)(10)(11)で $q = (x, y)$ に $m = (x_m, y_m)$ を代入して，次のように定まる．

$$V_l = \begin{pmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{pmatrix} = \begin{pmatrix} f_1(x_m, y_m) f_3(x_m, y_m) \\ f_2(x_m, y_m) f_3(x_m, y_m) \\ f_3(x_m, y_m) \end{pmatrix} \quad (12)$$

$$|V_l| = 1 \quad (13)$$

これを用いて、写真 P_a に関して視点 S_a からラベル l_a への方向ベクトルを V_{la} 、写真 P_b に関して視点 S_b からラベル l_b への方向ベクトルを V_{lb} とし、次のように記述する。

$$V_{la} = \text{Direction}(SP_a, l_a) \quad (14)$$

$$V_{lb} = \text{Direction}(SP_b, l_b) \quad (15)$$

空間スキーマ付き写真 P_a, P_b とラベル l_a, l_b へ方向 V_{la}, V_{lb} および POI が指し示す対象の位置 $M = (X_M, Y_M, Z_M)$ との関係は図 6 のようになる。

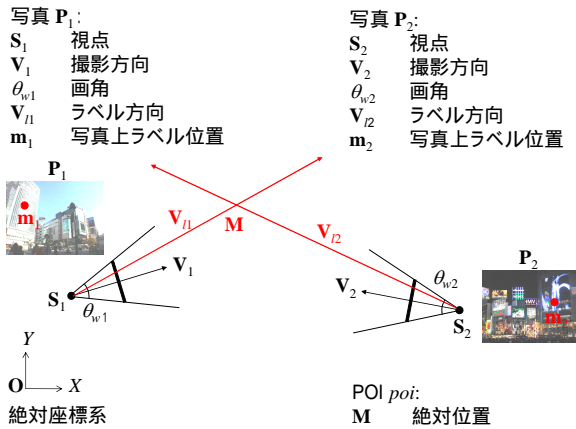


図 6. 2 つのラベルデータからの対象位置決定

図 6 に示すとおり、 M の座標は 2 つの半直線 $S_a + sV_{la}$ と $S_b + tV_{lb}$ (s, t は実数) との交点として算出されるが、これらが 3 次元空間上で完全に交わることは非現実的である。そこでこれらの半直線の XY 平面上への射影に関して求めた交点を (X_M, Y_M) とし、これらの半直線上で $(X, Y) = (X_M, Y_M)$ である点をそれぞれ M_a, M_b とすると、次式より (X_M, Y_M) が定まる。

$$M_a = S_a + sV_{la} \quad (16)$$

$$M_b = S_b + tV_{lb} \quad (17)$$

$$(X_M, Y_M) = (M_a.X, M_a.Y) = (M_b.X, M_b.Y) \quad (18)$$

Z_M は $M_a.Z$ と $M_b.Z$ との平均として算出する。

$$Z_M = \frac{M_a.Z + M_b.Z}{2} \quad (19)$$

以上より対象位置 $M = (X_M, Y_M, Z_M)$ が定まる。

次に、図 7 に示すとおり、同一の対象を指し示す 3 つ以上のラベルデータから POI データを作成する手順を示す。この場合、各写真の視点からラベル方向への複数の半直線は XY 平面上でも 1 点で交わらない。そこで 2 つラベルデータの組み合わせをすべて作成し、そのすべての組み合わせに関して提案手法により

$POI.M$ を算出し、それらの平均を最終的な $POI.M$ とする。以下に具体的に示す。同じ対象を指し示す n 個をラベルデータを $l_1, l_2, l_3, \dots, l_n$ とすると、これらがすべて同じ対象を指し示すため、

$$l_i.text = l_j.text \quad (1 \leq i, j \leq n, i \neq j) \quad (20)$$

である。このとき、

$$POI.text = l_i.text \quad (1 \leq i \leq n) \quad (21)$$

であり、 l_i, l_j ($1 \leq i, j \leq n, i \neq j$) から算出された POI の位置を M_{ij} とすると、次式が成立する。

$$POI.M = \sum_{1 \leq i, j \leq n, i \neq j} \frac{M_{ij}}{n C_2} \quad (22)$$

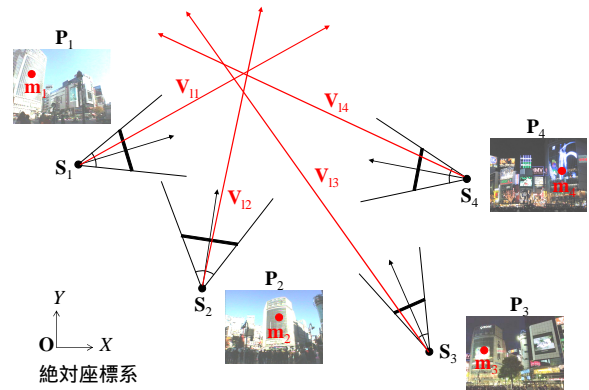


図 7.3 つ以上のラベルデータからの対象位置決定

4. 空間スキーマ付き写真へのラベルデータ伝播

ラベル共有システムにラベルデータを要求した特定の空間スキーマ付き写真に対して、POI 共有データベース中の POI データのうち、指し示す対象がその写真に写っているものすべてをラベルデータ化して伝播させる手法を示す。POI の共有データベースを DB_{POI} 、 DB_{POI} 中の POI データ集合を $POI_{DB_{POI}}$ 、ラベルデータを要求した写真データを $P_q = (I_q, SP_q, L_q)$ とする。 $L_q = l_{q1}, l_{q2}, l_{q3}, \dots$ である。目的は $POI_{DB_{POI}}$ から L_q を作成することであり、次のように段階的に処理する。

- (1) $POI_{DB_{POI}}$ から写真 P_q に対するラベルデータとなり得る POI データ集合 POI_q を作成する。
- (2) POI_q 中の POI データをラベルデータ化し、写真 P_q に対するラベルデータ集合 L_q とする。

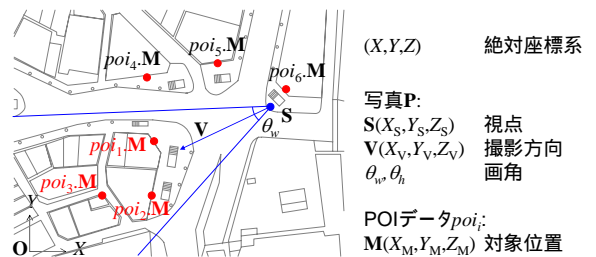


図 8. 空間スキーマ付き写真へ伝播する POI データ

例として、図 8 は空間スキーマ付き写真 P_q と POI データとの関係を XY 平面上で表したものであり、 S は P_q の視点、 V は P_q の撮影方向ベクトル、 θ_w は P_q の画角である。このとき poi_1 、 poi_2 が写真 P_q のラベルデータとなり得る POI データである。

上記(1)(2)の順に処理内容を具体的に説明する。InView を写真データ P の空間スキーマ SP と実世界座標上の点 A を引数とし、写真 P が切り取る空間内部に点 A が含まれるとき $true$ 、含まれないとき $false$ を返す関数とする。すなわち、

$$\text{InView}(SP, A) = \begin{cases} true & \text{点}A\text{が写真}P\text{の切り取る空間内} \\ false & \text{点}A\text{が写真}P\text{の切り取る空間外} \end{cases}$$

である。具体的には次の式 (26) を満たすとき $\text{InView}(SP, A) = true$ となる。

$$\left| \arccos \frac{(A - SP.S) \cdot SP.V}{\|A - SP.S\| \|SP.V\|} \right| \leq \frac{\theta_w}{2} \quad (26)$$

これを用いて P_q のラベルデータとなり得る POI データ集合 POI_q は次のように表される。

$$POI_q = \{ poi \mid poi \in POI_{DBPOI}, \text{InView}(SP, poi.M) = true \} \quad (27)$$

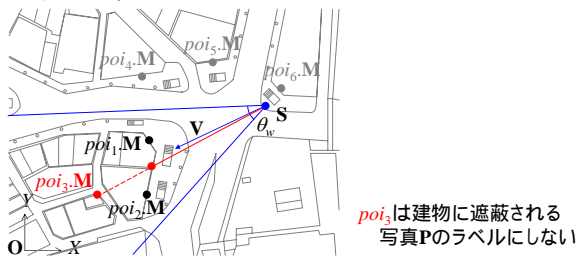


図 9. 地図データを利用した遮蔽の考慮

次に、写真 P_q の視点から見て建物等の裏側にある POI は、写真 P_q が切り取る空間内にあっても写真には写らないはずである。そこで、地図データを利用して建物等により遮蔽される POI データを POI_q から除外する。たとえば図 9 において、 poi_3 が建物に遮蔽されて写真 P_q に写らないため除外される。 POI_q から遮蔽されて見えない POI データを除外した集合 POI_q' は次のように表される。

$$POI_q' = \left\{ poi \mid \begin{array}{l} poi \in POI_q, \forall feature \in MAP, \\ \neg \text{Intersection}(feature, (poi.M - SP.S)) \end{array} \right\} \quad (28)$$

ここで MAP は地図データであり、対象地域に含まれる建物等の形状を表現する、点や点の列である線分や多角形からなる形状データ集合である。なお、本手法では地図データとして存在しないオブジェクトによる遮蔽は考慮されない。

続いて POI_q' 中の POI データをラベルデータ化し、ラベルデータ集合 L_q を作成する。ひとつの POI データからひとつのラベルデータを作成する。以降、ラベル

データ l から POI データ poi を作成する手順を述べる。ラベルテキストと関連 URL は POI の対応する要素をそのまま用いる。すなわち、

$$l.text = POI.text \quad (29)$$

$$l.url = POI.url \quad (30)$$

である。ラベルが指し示す対象の写真上での位置 m は、POI が指し示す対象の実世界での位置 M の、写真データ P_q に関する写真座標系への射影として算出する。いま関数 ProjectionToPhoto を写真データ P の空間スキーマ SP と実世界座標上の点 A を引数とし、写真 P の写真座標系上へ点 A を射影した点 a を返すものとする、ラベル l が指し示す対象の写真 P_q 上での位置 m は次の式で算出される。

$$m = \text{ProjectionToPhoto}(SP_q, M)$$

ここで関数 ProjectionToPhoto による m の算出方法を説明する。共線条件式(3)(4)は次のように変形できる。

$$x = -c \frac{a_1(X - X_s) + a_4(Y - Y_s) + a_7(Z - Z_s)}{a_7(X - X_s) + a_8(Y - Y_s) + a_9(Z - Z_s)} \quad (31)$$

$$y = -c \frac{a_2(X - X_s) + a_5(Y - Y_s) + a_8(Z - Z_s)}{a_7(X - X_s) + a_8(Y - Y_s) + a_9(Z - Z_s)} \quad (32)$$

式(31)(32)で $q = (x, y)$ に $m = (x_m, y_m)$ を、 $Q = (X, Y, Z)$ に $M = (X_M, Y_M, Z_M)$ を代入し、 M を既知として m が定まる。



図 10. ラベルデータの写真上への配置

最後に写真データ P_q の写真イメージ I_q 上にラベルデータ集合 L_q 中のすべてのラベルデータを配置した例を図 10 に示す。本章では、POI データの位置 M や空間スキーマ付き写真の空間スキーマ SP 等の実世界座標系上の値を、高さ Z 座標を含む 3 次元座標で議論してきたが、ここに Z 座標を持たない 2 次元のデータが含まれる場合は、写真上位置 m の x 座標のみ算出できるため、図 11 に示すように写真イメージの下部にラベルを配置する。



マツモトキヨシ
SuperLisa渋谷
Starbucks
Q Front
TSUTAYA

図 11. ラベルデータの写真下部への配置

5. プロトタイプシステム

提案手法をプロトタイプシステムとして実装した。最終的には不特定多数のクライアントを対象とし、多くて多数のユーザ間でラベルデータを共有する WEB 上のサーバシステムとして実装することを想定しているが、現時点ではローカルマシン上で動作し、そのマシン内の写真集合内でラベルデータを共有するシステムとして実装した。プロトタイプシステムにおいて、写真が持つメタデータ(空間スキーマ, ラベルデータ)をすべて Exif 情報として、写真ファイル内に記録した。本システムはユーザに以下の処理を提供する。

- (a) 特定の写真に対するラベルの取得
ユーザが選択した写真に対する複数のラベルの候補のリストが、ラベル共有データベースより作成され表示される。ユーザはリストから必要なラベルのみ複数選択し、写真に関連付ける。
- (b) 写真集合に対するラベルの一括取得
ユーザが選択したフォルダ中のすべての写真に対して、ラベル共有データベースよりそれぞれ複数のラベルが関連付けられる。
- (c) 写真上へのラベルの配置
ユーザが選択した写真が持つ複数のラベルが写真上に配置される。
- (d) WEB サイトのナビゲーション
(b)でラベルが URL を持つ場合、ユーザが写真上に配置されたラベルをクリックすると、WEB ブラウザが立ち上がり、WEB サイトが開かれる。
- (e) 新たなラベルの入力
ユーザが選択した写真上の 1 点をクリックし、ラベルテキストと URL (任意) を入力することで、新たなラベルデータが作成され、その写真に関連付けられる。
- (f) 入力したラベルの共有データベースへの登録

ユーザが共有を許可したラベルをラベル共有データベースに登録する。登録したラベルと同一のラベルテキストを持つラベルデータが共有データベース中に存在した場合、そのラベルデータの対象位置が決定して POI データ化され、(a)(b)の処理により他の写真が取得することが可能となる。

- (g) ラベルテキストによる写真の検索
ラベル共有リスト中の全ラベルのリストから、ユーザによって選択されたラベルとして持つ写真のリストを検索する。

図 12 は、プロトタイプシステムにおいて赤色で表示されたラベル”GFront”が赤枠で囲まれた 2 枚の写真上でユーザにより入力されて POI データ化し、他の複数の写真に伝播した様子を示している。

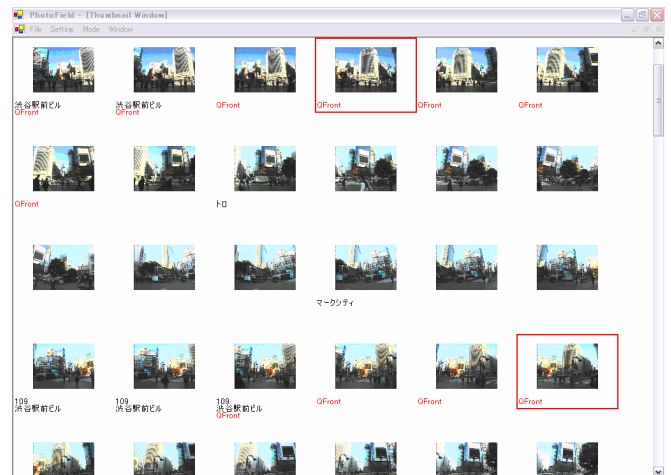


図 12. プロトタイプシステム

6. 結論と今後の課題

本論文では、今後空間メタデータを持つ写真(空間情報付き写真)が WEB 上で流通することを想定し、ラベルデータを POI データ化することで複数の写真間で共有する手法を提案し、プロトタイプシステムとして実装した。提案手法は次のような特徴および利点を持つ。

- ・ ラベルデータを POI データ化することで写真データから独立させ、その再利用性を高める。
- ・ 複数の空間情報付き写真上で同じラベルが入力されると、対象の実世界での位置が定まりラベルデータが POI データ化する。すなわち、多くのユーザにより入力されたラベルが POI データとなり共有される。
- ・ 空間情報付き写真に対して、その写真に写っているすべての POI データがラベルデータとして伝播する。
- ・ 伝播した複数のラベルテキストが写真上の

適切な位置に配置される。ラベルデータが関連 URL を持っている場合、写真がクリップ可能なものとなる。

- ・ 地図データ等として整備されている既存の POI データを空間情報付き写真のラベルデータとして利用できる。
- ・ 作成した POI データは地図データとして他のアプリケーションで利用できる。

逆に提案手法は次のような欠点を今後の課題として持つ。

- (a) 空間情報付き写真のみを対象としているが、現時点ではそのような空間メタデータを持たない写真のほうが多く、GPS 等の空間センサの精度は低いため空間メタデータを持っていてもその精度は低い。そのため提案手法を適用できるデータが少ない。
- (b) 1 枚の空間情報付き写真が切り取る空間 (FOV) はその撮影方向に無限に広がっているため、4 章で述べたとおり FOV に含まれるすべてのラベルデータをその写真に伝播させると、大量のラベルデータが伝播し、その多くが建物等に遮蔽されているか、距離が遠すぎて見えない対象に対するラベルとなる。

(a) に関して、本研究は測位技術の進歩により数 m 精度の高精度の空間メタデータを GPS 携帯電話等によって安価に取得できる環境を想定しているが、現状はまだそのような環境にならないため、[3]において、空間メタデータを持たない通常の写真を対象とし、ユーザが地図上および写真上の対応する 3 点を入力することで空間情報付き写真が前提とする空間メタデータを作成するツールを提案しており、プロトタイプシステムで用いた空間情報付き写真データ作成に利用している。

(b) に関して、視線の長さに閾値を決定することで空間情報付き写真が切り取る画角内の空間 (FOV) を有限にし、この FOV 内のラベルデータを伝播させるという手法が考えられる。これはコンピュータグラフィクスにおけるクリッピング処理にあたる。ここで、視線の長さを決定することは、写真に写る空間の奥行きを定めることにあたり、すべての写真に対して一定の閾値を定めることは不自然である。たとえば、丸ビルからすぐ隣の東京駅を写した写真と、同じく丸ビルから富士山を写した写真では、写真に写る空間の奥行きはまったく異なるためである。そこで、視線と地図データとの交点や、

Exif 情報の対象距離や対象距離範囲等を用いて写真ごとに視線の長さの閾値を決定する手法を検討する予定である。また、写真ごとにその写真に写る空間の範囲を決定する手法とは逆に、ラベルごとにその対象が見える領域 (ラベルの有効領域) を決定できる可能性もある。この場合は、ラベルの有効領域と FOV が重なる写真に対してラベルを伝播させる。

文 献

- [1] Digital Still Camera Image File Format Standard (Exchangeable image file format for Digital Still Cameras: Exif), JEIDA (Japan Electronic Industry Development Association), 1988.
- [2] 藤田秀之, 有川正俊, 岡村耕二, “高精度な空間情報付き写真の 3 次元実空間マッピング,” 電子情報通信学会論文誌, Vol.J87-A, No.1, pp.120-131, 2004.
- [3] 藤田秀之, 有川正俊, “写真データの空間的な組織化と視覚化,” 電子情報通信学会データ工学研究会, 第 14 回データ工学ワークショップ DEWS2004, 2004.
- [4] “Google Image Search”, <<http://images.google.com>>, 2001.
- [5] Mor Naaman, Andreas Paepcke, Hector Garcia-Molina, “From Where To What: Metadata Sharing for Digital Photographs with Geographic Coordinates,” 10th International Conference on Cooperative Information Systems (COOPIS), 2003.
- [6] “PhotoShop Album,” <<http://www.adobe.co.jp/products/photoshopalbum/overview/>>, 2003.
- [7] “G-XML と POI と辞書,” データベース振興センター, G-XML 講演会予稿集, 2004.
- [8] 村井俊治, “空間情報工学,” 日本測量協会, pp.175, 2001.