

位置・姿勢情報付き写真検索アルゴリズムの検討

土斐崎 博[†], 安藤 嵩[†], 金井 敦[†]

[†]法政大学大学院 工学研究科

あらまし

近年、GPSや電子コンパス、加速度センサを搭載した携帯電話やカメラが普及している。その携帯電話のカメラ機能を使えば、写真に位置情報や姿勢情報をメタデータとして自動的に記録することが可能になる。今後このようなメタデータを付加した写真が普及すれば、膨大な写真データベースから、自分の見たい写真をより的確に検索することが可能になる。本稿では、撮影位置の情報に、カメラの撮影方位、仰角、回転角、画角を新たに加えることで、見たい場所、見たい対象物の写っているなどの最もユーザーの嗜好に合った写真を検索する手法を提案する。さらに、実際に評価を行う。

A Study on an Algorithm to Search Photograph with Location and Attitude Information

HIROSHI DOIZAKI[†], TAKASHI ANDO[†], ATSUSHI KANAI[†],

[†]Graduate School of Engineering, Hosei University

Abstract

In recent years, mobile telephone equipped with GPS, electronic compasses and acceleration sensor have been familiar. We can record location and attitude information into picture as metadata automatically when we use the above functions. We will be able to more exactly search photographs we need if the photograph with such a meta data circulate in future. In this paper, we propose a methodology to search photographs which contain the desired object by adding direction, elevation, roll, and angle as metadata. And the methodology is evaluated using actual photographs.

1. はじめに

近年、デジタルカメラや携帯電話の普及により、写真を撮るという行為が非常に多くなってきた。また、GPSや電子コンパス、加速度センサを搭載した携帯電話が普及し始めている。その携帯電話のカメラ機能を使うと、写真にGPS情報と方位の情報を自動的にメタデータとして記録することができる。

今後、このような撮影位置情報と方位情報のみならず、どの角度で撮ったのかといった撮影姿勢情報をもった写真を撮影できるデジタル機器が普及し、そういった情報をもった写真がWeb上で流通することが考えられる。本稿ではこのような、撮影位置情報と撮影姿勢情報をもった写真を「位置・姿勢情報付き写真」と呼ぶ。つまり、どこから、どの方向に、どれくらい角度で、どれ位の大ききで撮影したのかという情報が付加された写真である。この「位置・姿勢情報付き写真」を用いれば、写真を検索するうえでより高度な検索が可能となる。

現状では、写真を撮影位置から検索するサービスが存在する。たとえば、Google Mapで地図上に各々の撮影場所の写真を表示させる機能等がそれにあたる。つまり、どこから撮影した写真なのかという検索のみであり、写真に写っているものまでを検索することができなかった。しかし、「位置・姿勢情報付き写真」を用いることで、何を写している写真なのかを、計算処理することである程度想定することが可能になる。

そこで、本稿では「位置・姿勢情報付き写真」を用いて、ユーザーの目的に合致した対象物が映る写真を優先順位を付けて選び出すアルゴリズムを提案する。そして、そのアルゴリズムを実際に実装し、評価する。

2. 仮定する画像情報

本稿では、どの位置から、どの方向に向かって、どれ位の画角で撮影したのかという情報を付加した写真を「位置・姿勢情報付き写真」と呼ぶ。この写真の情報には、具体的には、緯度、経度、高度、仰角、方位角、回転角、カメラの画角の情報がある。緯度、経度、高度の情報はGPS機器から取得し、方位角は電子コンパス、仰角、回転角はジャイロセンサーから取得し、画角はカメラのレンズから算出する。そして、取得した情報は、写真のExif Format上に載せる。表1に

仕様をまとめる。

3. 検索手法の提案

3.1 目標物とカメラのモデリング

対象物を特定するために、対象物の中心を点とし、その対象物全体を球で指定する。対象物を球で捉えたのは、様々な方向からの撮影に対しても柔軟に対処でき、容易に計算ができるからである。

まず、対象物の中心点を、地球上のある場所の点で指定する。そのため、対象物の緯度、経度、高度の情報が必要である。さらに、対象物全体を球として捉えるため、その球の半径の情報も必要である。ゆえに、本稿では、対象物を特定するには、対象物の緯度、経度、高度、半径(大きさ)の情報を用いることとする。また、カメラの視野に関しては、計算を容易にする為に、円で想定することとする。

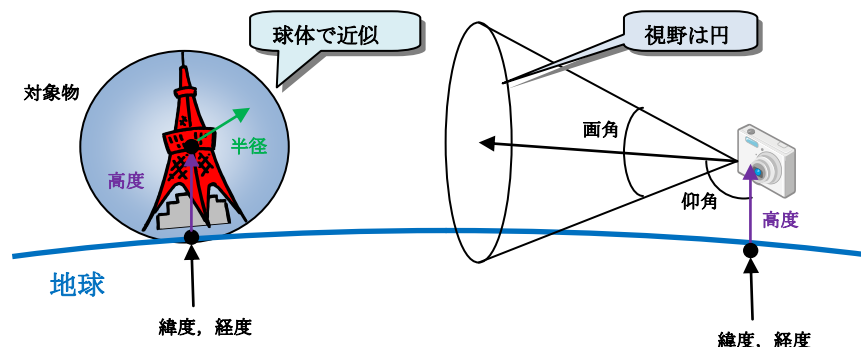


図1 目標物とカメラのモデリング

3.2 最適な写真を判断する為の三要素

対象物が写っている写真を選択するうえで、基準となる数値を定める必要がある。

そこで、どの写真に対象物が最適に写っているかを判断する為に、「中心偏差 (U)」、「対象物全形率 (M)」、「対象物占有率 (N)」の3つの値を定義する。この3つの値はカメラの視野と対象物との関係を表した数値であり、対象物が最適に写っている写真を導出するのに用いる。

以下に定義を示す。

「中心偏差 (U)」は、カメラの視野の中心と対象物の中心点がどれだけずれているかを表す値として定義する。

「対象物全形率 (M)」は、対象物の全形がどれだけの割合でカメラの視野に写っているかを表す値として定義する。

「対象物占有率 (N)」は、対象物がカメラの視野にどれだけの大きさで写っているかを表す値として定義する。

次節で、各々の詳細を説明する。

3.2.1 中心偏差

カメラの視野の中心と対象物の中心点がどれだけずれているかの値を「中心偏差」と定義する。この値が1の場合対象物の中心点が視野の中心にあり、小さくなれば視野の外側に近づいていき、0なった場合は視野内から外れていること

表1 位置・姿勢情報付き写真の仕様

写真情報	詳細	取得機器
位置情報	緯度	GPS 機器
	経度	
	高度	
姿勢情報	方位角	電子コンパス
	仰角	ジャイロセンサー
	回転角	
カメラ情報	画角	カメラのレンズから算出

とする。つまり、中心偏差は1に近ければ近い程、対象物を中心に写している写真であると判断できる値である。

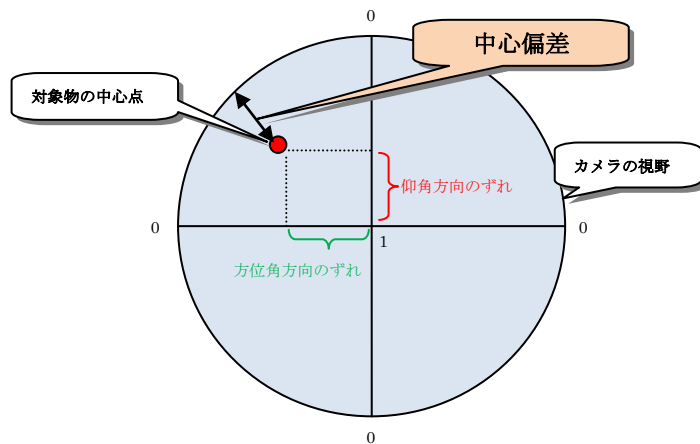


図2 中心偏差

中心偏差 U を求める。図2参照の仰角方向のずれ (P) と方位角方向のずれ (E) から、カメラの視野外と対象物の中心点との距離を算出すると、以下のようになる。

$$U = 1 - \sqrt{P^2 + E^2} \quad (1)$$

但し $U \leq 0$ の場合 $U = 0$ とする

3. 2. 2 対象物全形率

対象物の全形がどれだけの割合でカメラの視野に写っているかの値を「対象物全形率」と定義する。例えば、対象物の全形を写している写真であれば、対象物全形率は1となり、対象物を半分程度しか写していなければ、0.5となり、全く写していなければ0になる。つまり、対象物全形率が1に近ければ近い程、対象物の全形をより多く写している写真であると判断できる値である。

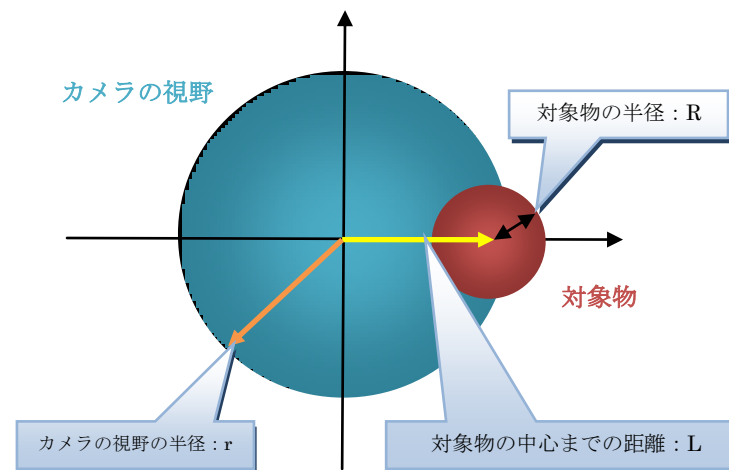


図3 カメラの視野と対象物

カメラの視野と目標物が重なる部分の長さを対象物の直径で割ることにより求めることとする。上記図3より、対象物の半径を R 、カメラの視野の半径を r 、対象物の中心までの距離を L とする。 R と r の大きさの関係で、対象物全形率 M の求め方は変わってくる。そこで、 $r \geq R$ の場合と、 $r < R$ の場合の2通りの場合分けを行う必要がある。さらに、カメラの視野と対象物の位置の関係でも3通りの場合分けをする必要がある。以下に式を記す。

$r \geq R$ で且つ、

$$\begin{cases} 0 \leq L \leq r - R \text{ のとき} & M = 1 \\ r - R < L < r + R \text{ のとき} & M = \frac{r+R-L}{2R} \\ r + R \leq L \text{ のとき} & M = 0 \end{cases} \quad (2)$$

$r < R$ で且つ、

$$\begin{cases} 0 \leq L \leq R - r \text{ のとき} & M = \frac{r}{R} \\ R - r < L < r + R \text{ のとき} & M = \frac{r+R-L}{2R} \\ r + R \leq L \text{ のとき} & M = 0 \end{cases} \quad (3)$$

3. 2. 3 対象物占有率

カメラの視野に対象物がどれ位の大きさで写っているかを表した値を「対象物占有率」と定義する。カメラの視野全体に対象物が写っていれば1とし、対象物の全形を捉えていても、カメラの視野全体の半分の大きさしか写っていなければ0.5、カメラの視野に対象物が全く写っていなければ0とする。つまり、対象物占有率が1に近ければ近い程、対象物を大きく写している写真であると判断できる値である。

前節と同様に、カメラの視野と目標物が重なる部分の長さを対象物の直径で割ることにより求めることとする。

R と r の大きさの関係で、対象物占有率 N の求め方は変わってくる。そこで、 $r \geq R$ の場合と、 $r < R$ の場合の2通りの場合分けを行う必要がある。さらに、カメラの視野と対象物の位置の関係でも3通りの場合分けをする必要がある。以下に式を記す。

$r \geq R$ で且つ、

$$\begin{cases} 0 \leq L \leq r - R \text{ のとき} & N = \frac{R}{r} \\ r - R < L < r + R \text{ のとき} & N = \frac{r+R-L}{2r} \\ r + R \leq L \text{ のとき} & N = 0 \end{cases} \quad (4)$$

$r < R$ で且つ、

$$\begin{cases} 0 \leq L \leq R - r \text{ のとき} & N = 1 \\ R - r < L < r + R \text{ のとき} & N = \frac{r+R-L}{2r} \\ r + R \leq L \text{ のとき} & N = 0 \end{cases} \quad (5)$$

3. 3 評価関数

前節で定義した3つの値を用いて、どの写真が対象物を最も良く写しているかを判断する評価関数を定義する。

対象物が写っている写真を見たい人にとって重要な問題は以下の3つである。

- a. 対象物の全形が写っている
- b. 対象物が大きく写っている
- c. 写真の中心に対象物が写っている

上記の3つの項目は、 a が対象物全形率、 b が対象物占有率、 c が中心偏差と関係している。

つまり、一番欲しい写真は「対象物の全形が大きく中央に写っている写真」である。例えば、対象物の全形が写っているが、写真全体には小さく写っている場合や、写真全体に対象物が写っているが、対象物の一部分しか写っていない写真は上記の基準から判断して好ましくない。

上記の考えに従って、評価関数 S を以下のように表す。

$$S = \beta U + (1 - \beta)(MN) \quad (6)$$

上記の式で、カメラの視野に対象物の全形がどれ位写っているかの割合 M と視野に対象物が大きく写っているかの割合 N を掛け合わせることで、「対象物の全形が大きく写っている写真」であるかということの評価している。

また、対象物が中央に写っているという評価に関しては、ユーザーの嗜好に最も影響を受けるものと判断し、重みづけの数値 β をつける。

なお、評価関数 S の値の最大値は1であり、より「対象物の全形が大きく中央に写っている写真」であることを意味し、0であれば、対象物が全く写っていない写真を意味する。

3. 4 全体の計算アルゴリズム

本節では、前節で述べた評価関数を求める計算アルゴリズムを示す。以下に示す手順が、位置・姿勢情報付き写真から評価関数を求める概括的な流れである。なお、具体的な計算式に関しては本稿では省略する。

【手順1】写真の位置情報と対象物の二点間の直線距離を求める。

【手順2】撮影位置から対象物が地平線の影響で物理的に見ることが不可能な写真であるかを判断する。

【手順3】撮影位置からみた、対象物への二点間の仰角を求める。

【手順4】撮影位置からみた、対象物への二点間の方位を求める。

【手順5】中心偏差を求める。

【手順6】対象物全形率を求める。

【手順7】対象物占有率を求める。

【手順8】評価関数を求める。

手順1は、地球の丸みを考慮しない2点間の直線距離を求める。そのために撮影位置と対象位置の2点を地球の核を中心とした球座標に変換し距離を算出した。

手順2は、撮影位置と対象位置の二点間の直線と地球の核との距離を求め、地球の半径より小さければ、対象物は物理的に見ることはできないと判断する。

手順3では、撮影位置、対象物、地球の核の3点を結んだ三角形から撮影位置の角度を求める。

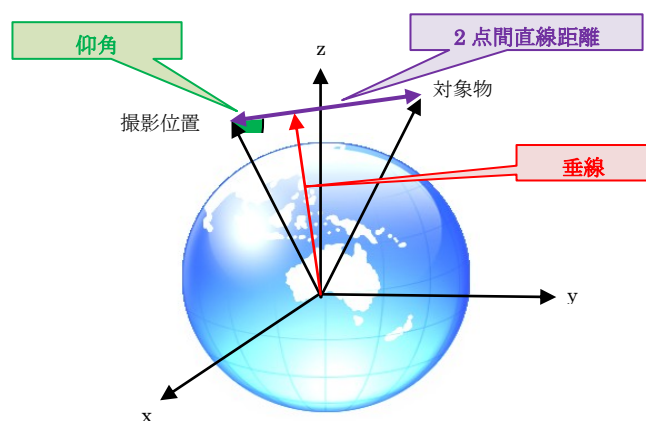


図4 求める数値

手順4は、参考文献[3]を基に求める。

手順5, 6, 7, 8は前節に詳細を述べた。

4. 評価

本論文で示した手法の有効性を確かめるため、前章で述べたアルゴリズムを実装し、検証した。サンプルデータとして、Webカメラで撮影した写真にGPSや電子コンパス、ジャイロセンサなどの機器を用いて、位置・姿勢情報付き写真を撮影する。大学近辺にあるランドマークである「スカイツリー西東京」を対象物とし、撮影位置や写真の構図を様々に変更しながら70枚撮影した。なお、今回

撮影機器として使用した電子コンパスに関しては十分な精度を得られなかったため方位情報に補正を加えた。また、カメラの画角に関してはズーム機能の持たないWebカメラを使用したため、どのサンプル写真も画角は一定である。

各々の位置・姿勢情報付き写真に、対象物を「スカイツリー西東京」に指定した検索アルゴリズムで評価関数 α を求め、その評価関数の高い順に並べ替える。重みづけの数値 β の値を0.1, 0.5, 1.0と変化させて検索した。

以下に $\beta = 0.5$ の場合の上位5枚の写真、真ん中(33位~37位)5枚の写真、下位5枚の写真を図5に示す。



図5 $\beta = 0.5$ の検索結果

上位5枚の写真(図5)をそれぞれ比較してみると、どの写真にも対象物が写っていることが分かる。また、真ん中5枚の写真には、やや小さめに中心よりずれている位置に対象物が写っている写真が多く見受けられる。そして、下位5枚の写真に関しては、全く対象物が写っていない写真が検索された。このことから、より対象物の全形を大きく写している写真が順序よく検索されており、評価関数がほぼ有効であることがわかる。

次に以下に各々の β の値の上位 5 枚の写真を図 6 に示す。



図 6 β が各々の場合の上位 5 枚の検索結果

β の値を各々変化させた上位 5 枚の写真 (図 6) を比較してみる。 $\beta = 0.5$ の上位 5 枚の写真と $\beta = 1.0$ の上位 5 枚の写真に関してはほぼ同等の写真が上位に検索されている。しかし、 $\beta = 0.1$ の上位 5 枚の写真と $\beta = 0.5$ の上位 5 枚の写真とを比較すると、 $\beta = 0.1$ の写真は対象物が大きく写っているものが比較的多く、 $\beta = 0.5$ の写真は対象物が中央に写っているものが多いことがわかる。このことから、対象物が中央に写っているという重みの数値 β が確実に機能していることがわかる。

5. おわりに

本稿では「位置・姿勢情報付き写真」を用いて、ユーザーの目的に合致した被写体が映る写真を選び出すアルゴリズムを提案した。また、優先順位を決める為に各々の写真の対象物の写り具合を判断する評価関数を定義した。そして、サンプル写真を大量に撮影し、各々の写真の評価するシステムを実装し、提案した評価

手法の有効性を検証した。その結果、評価関数の上位にあたる写真に関しては、多少の誤差はあるもののほぼユーザーの目的に合致した被写体が写った写真を検索できることがわかった。

今後の課題としては、撮影機器の影響で検証の対象に入らなかったカメラの画角が変化した場合でも、提案した評価手法が有効であるかを検証する必要がある。また、仰角方向においてサンプル写真と理論値との誤差が多少大きい原因の追及、複雑な計算処理の簡略化等が挙げられる。

参考文献

- [1] 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, A, Vol. J87-A, No. 1, pp. 120-131(Janury2004)
- [2] 鈴木雄大, “撮影位置・姿勢情報に基づく観光地映像ブラウジングシステム”,
奈良先端科学技術大学院
- [3] Survey Review, 33, 261(July 1996) ”TOTAL INVERSE SOLUTIONS FOR THE GEODESIC AND GREAT ELLIPTIC”
- [4] 国土交通省国土地理院 距離と方位計算,
<http://vldb.gsi.go.jp/sokuchi/surveycalc/bl2stf.html>