Vol.2010-DBS-151 No.26
2010/11/13

意味的アフィン変換:重ね合わせ操作による直感的なコンテンツ融合モデル

武田　十季[†1]　鶴野　玲治[†2]　牛尼　剛聡[†2]

日常空間には様々な情報表現（コンテンツ）が存在し，人はそれらを利用して情報活動を行う．そうした際，複数のコンテンツに含まれる情報を統合して利用することが多い．本研究では，実空間上に存在する複数のコンテンツを，カメラ付き携帯電話を利用して重ね合わせる操作によって，相互の意味内容に基づいて仮想的に情報を融合し提示する Pick and Lap 手法を提案する．本手法を実現するために，コンテンツの仮想的な融合の一般的なモデルとして，意味的アフィン変換を提案する．2 次元アフィン変換は，2 次元平面上の平行移動，回転，拡大縮小，剪断を表現するために利用される幾何変換であるが，意味的アフィン変換では，コンテンツの重ね合わせの対象となる 2 つのコンテンツの意味内容に基づいてアフィン変換のパラメータを決定することで，コンテンツが融合可能となる．

Semantic Affine Transformation: an Institutive Information Integration Model By Content Overlapping

Toki Takeda ,[†1] Reiji Tsuruno [†2] and Taketoshi Ushiama [†2]

Various presentations of information (contents) exist in our daily life, and we use them for activities.In this case, we often integrate the information included in multiple contents.In this paper, we propose the *"Pick-and-Lap" technique* that enables users to virtually integrate information based on the semantic relationships of the contents, by means of overlapping one content with another using a camera phone.To develop systems based on the Pick-and-Lap technique , we introduce *Semantic Affine Transform* as a general model for integrating contents virtually.The affine transform is a popular geometric transform to express the parallel transfer, rotation, scaling and shearing.The semantic affine transform allows to integrate two overlapped contents based on the parameters of affine transform that are decided according to the semantic relationships among the contents.

1. はじめに

情報技術および情報化社会の発展に伴い，日常の活動の中で必要な情報を取得する要求が高まっている．それらの要求を満たすために，携帯端末を利用してユーザのニーズに合った情報を提示する研究が盛んに行われている．例えばオンライン地図をはじめとする地図サービスでは，現在地の地図の表示や目的地までのルート表示，また乗り換え案内など携帯端末に向けて機能が多様化され，さらに利便性が増してきている[1)2)3)]．しかし，デジタル・コンテンツによる情報の閲覧が普及しているにも関わらず，人はカレンダーや地図などの実空間上で実体を有する情報表現（以下，実空間コンテンツと呼ぶ）を用いて必要な情報を取得することがある．

人が実空間コンテンツから情報を取得する際には, 複数の実空間コンテンツを相互に参照して取得することが多い．例えば，店舗紹介が掲載された情報誌では，店舗の一覧ページと地図のページが存在し，それらコンテンツを相互に比較することによって，店舗の場所を地図で確認することがある．また，道端のポスターで鑑賞したい映画を見つけた際には，ポスターや映画の上映スケジュールと自分のスケジュール帳を用いて，日時の確認を行うこともある．このように，私たちは日常の中で，複数の実空間コンテンツを相互に参照し，情報を融合することで，必要な情報を獲得している．

このように，私たちは日常の中で，複数の実空間コンテンツを相互に参照し，情報を融合することで必要な情報を獲得しているが，一方で実空間コンテンツは，可変的でなく新たに情報を付加することが出来ず，融合された情報は可視化されないため，ユーザにとって負担となることがある．そこで本研究では携帯端末を用いて，参照する組み合わせによって情報を融合し，ユーザへ提示することによって，より効果的な情報閲覧への実現を目指す．

実空間コンテンツを拡張して, 付加的な情報を提示する研究に拡張現実 (AR) がある[4)]．拡張現実では端末上のディスプレイで，空間や実空間コンテンツに対して仮想的に情報を重

†1 九州大学大学院芸術工学府芸術工学専攻
Graduate School of Design, Kyushu University
†2 九州大学大学院芸術工学研究院
Faculty of Design, Kyushu University

© 2010 Information Processing Society of Japan

畳して提示することによって付加的な情報を可視化している．しかし，従来の拡張現実では，付加的に提示される情報は，空間や実空間コンテンツにあらかじめ対応づけられた情報に限られており，異なる複数の実空間コンテンツの関係性に着目した研究は行われていない．

本研究では，AR 技術を利用して，携帯電話上で実空間上の複数コンテンツを重ね合わせる操作を行うことで，コンテンツを意味的な関連に基づいて仮想的に融合することにより，ユーザの情報活動を支援する Pick and Lap 操作を提案する．本手法では， 実空間上のコンテンツを， 携帯電話のカメラを用いて拾い上げ (Pick)，実空間上の他のコンテンツへ重ね合わせる (Lap)． このとき，Pick したコンテンツと Lap するコンテンツの意味的な関係に基づき，2 つのコンテンツを仮想的に融合し，ユーザの目的に合致した形式で情報を提示する．

本手法を実現するためには，様々な技術的な課題を解決しなければならない．それらを大別すると次のようになる．

(1) 携帯電話のカメラを利用した実空間コンテンツの認識手法の開発
(2) 重ね合わせによる実空間コンテンツの仮想的な融合の論理的なモデルの開発
(3) サービス提供のためのインターネットを介したメタデータの管理機構の開発

本論文では，特に重ね合わせによる実空間コンテンツの仮想的な融合の形式的なモデルの開発に焦点を当てる．コンテンツを仮想的に融合するために，実空間コンテンツのレイアウトを考慮した情報の融合を行う形式的なモデルとして意味的アフィン変換を提案する．

2. Pick and Lap に基づく情報融合

2.1 概　　要

本研究の基盤となる Pick and Lap 操作に基づくコンテンツ融合の概要を述べる．Pick 操作とは， カメラ付き携帯電話で実空間上のコンテンツを撮影することにより， 対象とするコンテンツ全体，またはコンテンツの一部を選択する操作である． ここで，選択されたコンテンツ全体またはその一部を「Pick されたコンテンツ」と呼ぶ．一方， Lap 操作は，Pick されたコンテンツを， 他の実空間上のコンテンツ全体，またはコンテンツの一部分へ重ね合わせる操作である．ここで，Pick されたコンテンツを重ね合わせる対象となるコンテンツを「Lap されたコンテンツ」と呼ぶ．

Pick 操作と Lap 操作を実現するためには，対象とするコンテンツの全体，またはコンテンツの一部分を認識する必要がある．本研究では，コンテンツの認識対象として， 2 次元の矩形画像を利用する．また，Pick and Lap 操作が行われた際のコンテンツの融合を， 2 次元矩形画像の移動，拡大・縮小および画像の重ね合わせによって実現する．

Pick andLap 操作が行われると， 相互のコンテンツの意味的な関係性に基づいて，Lap される側のコンテンツが持つ意味的な空間が決定され，Pick されたコンテンツはその意味的空間上での適切な位置が定められる． つまり， Pick されたコンテンツと Lap されるコンテンツに含まれる情報が AR 技術を利用して融合されユーザに提示される．

本研究では，対象とする実空間上のコンテンツの例として，旅行情報誌を利用する (図 1)．旅行情報誌には， 観光スポット，ホテル，店舗， 土産に関する情報を紹介するページや，それらの場所を確認するための地図のページ等， 様々な種類のコンテンツが含まれている．本研究では，観光スポット，レストラン，ホテル等に関する個々の紹介記事および個々の地図や時刻表等を Pick 操作および Lap 操作の単位とする．ここで，Pick 操作や Lap 操作の対象となるコンテンツの部分を「コンテンツ・セグメント」と呼ぶ．また，携帯電話のカメラを利用して Pick 操作または Lap 操作の対象となるコンテンツセグメントをの認識するために， 地図やそれぞれの紹介記事に付加されている矩形画像を用いる (図 2)．

旅行情報誌を対象にしたコンテンツの重ね合わせによる具体的な表現は 3 章の融合モデルで述べる．

2.2 処理の流れ

Pick 操作，または Lap 操作を行うと，実空間コンテンツに含まれるコンテンツ・セグメントは携帯端末のカメラを利用して矩形画像として認識される．システムはあらかじめそれぞれのコンテンツ・セグメントに対応する矩形画像が登録されており，画像の特徴量に基づいて対象とするコンテンツ・セグメントが同定され，コンテンツ・セグメントに対応するメタデータが取得される (図 3)．

メタデータには店舗や地図などコンテンツ・セグメントの種類が含まれ，さらに，それぞれのコンテンツ・セグメントが表現する意味的なデータが含まれる．ここで，意味的なデータとは，たとえば，地図の場合では対象領域，店舗の場合では店舗の位置を示す緯度経度，開店時間，座席数などの情報が含まれる．

Pick されるコンテンツ・セグメントと Lap されるコンテンツ・セグメントの種類の組み合わせや順番に基づいて，異なる形式でコンテンツが融合され，ユーザに提示される．本論文では，コンテンツの融合方式をコンテンツ融合モデルとして形式化する．コンテンツ融合モデルは，ユーザに対して自然と感じさせる提示方法のことであり，詳細は次章で述べる．

コンテンツ融合モデルでは，対象とするコンテンツ・セグメントのメタデータに含まれる意味的なデータに基づいて，コンテンツ・セグメントを融合するためのパラメータを決定す

© 2010 Information Processing Society of Japan

Vol.2010-DBS-151 No.26
2010/11/13

図 1 Pick and Lap 操作

図 2 コンテンツ・セグメントと矩形画像

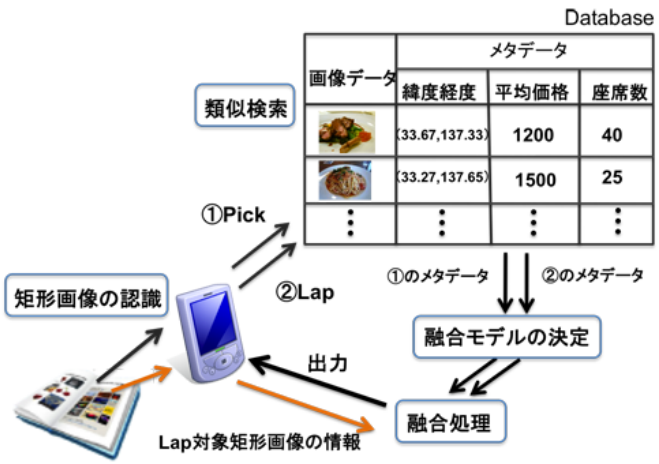

図 3 処理の流れ

る．このとき，携帯端末の画面上で Lap されるコンテンツ・セグメントに合わせて重畳するために，カメラから認識され画面に表示される Lap 対象コンテンツの矩形画像の大きさを取得して利用する．本論文では，この融合モデルを表現する式を形式化するために意味的アフィン変換を提案する．

3. コンテンツの融合モデルの決定

複数のコンテンツ・セグメントを重ね合わせる組み合わせと順番には，多数のパターンが存在する．重ね合わせのパターンによって，ユーザが要求する情報に適したコンテンツ融合モデルを用いて，コンテンツ・セグメントを融合してユーザに提示する．

ここでは，旅行情報誌を例として，コンテンツ融合モデルについて説明する．旅行情報誌には様々なコンテンツ・セグメントが含まれているが，ここでは店舗紹介記事 (店舗) と地図，観光スポットの 3 種類のコンテンツ・セグメントを対象にして以下の 3 つの重ね合わせのパターンに対する表現方法を述べる．

(1) 「店舗」を「地図」へ重ね合わせる
(2) 「地図」を「店舗」へ重ね合わせる
(3) 「店舗」を「観光スポット」へ重ね合わせる

「店舗」を「地図」へ重ね合わせた場合，携帯端末の画面上では「店舗」の矩形画像を地図上で「店舗」が存在する位置に合わせて表示する（図 4:左 ）．これによって，ユーザは興味を持った「店舗」を，閲覧したい「地図」上へ重ね合わせることで，店舗の位置を確認することができる．また，地図には同じ地域を示す地図でも略地図などのように種類の異なる地図が存在するが，ユーザは重ね合わせる操作によって，どの地図上でも店舗の位置を確認することが可能となる．さらに複数の店舗を Pick して地図へ Lap すると，一度にそれぞれの店舗の場所を確認することも可能となり，従来の旅行情報誌で閲覧するときのように店舗一覧ページと地図を交互に複数回参照する必要性がなくなる．

「地図」を「店舗」へ重ね合わせた場合，ユーザが Pick した「地図」上で「店舗」の場所がどこに存在するかを確認できるように,「店舗」の矩形画像領域内に「店舗」の場所を中心とした「地図」を表示する（図 4:中央 ）．つまりユーザは，ある特定の地域を示した地図や，縮尺の異なる地図など，様々な地図の中から閲覧したい地図を選んで Pick し，興味のある店舗へ Lap することで，Pick した対象の地図上においてそれぞれの店舗がどのあたりに存在するかを確認可能となる．

「店舗」と「観光スポット」を重ね合わせた場合,「店舗」と「観光スポット」の位置関係の比較を行えるように提示する．例えば，ユーザが特定の観光スポットに行こうと思った際に，その観光スポットから近い飲食店で食事したいと考えたとする．このような場合に，ユーザは興味を持った複数の「店舗」を選んで Pick し,「観光スポット」へ Lap すると，それぞれの店舗が観光スポットから，どのくらい離れ，どの方角に位置するかということを確

© 2010 Information Processing Society of Japan

Vol.2010-DBS-151 No.26
2010/11/13

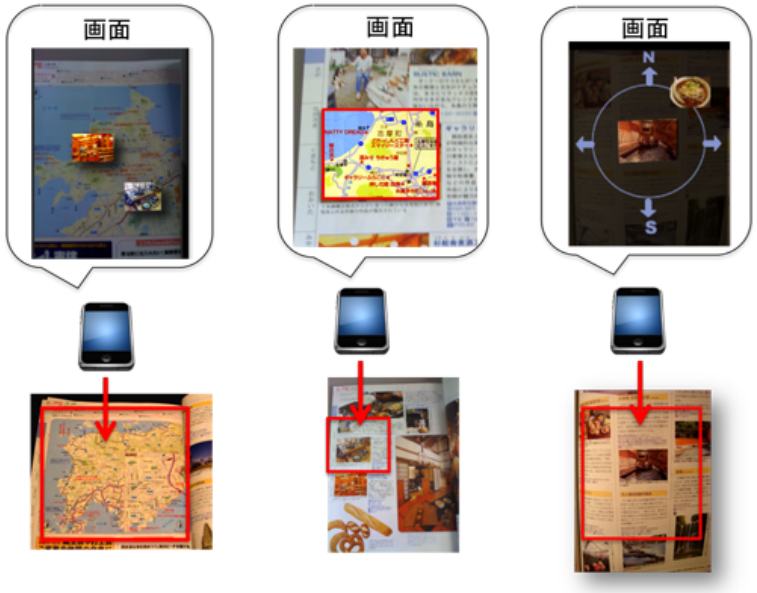

図 4 コンテンツ融合モデル 左:「店舗」を「地図」に 中央:「地図」を「店舗」に 右:「店舗」を「観光スポット」に

認できる．

位置関係を示すために，図 4 :右のように,「観光スポット」を中心にした方角を示す軸を仮想的に作ることで「観光スポット」から見た「店舗」の方角を示す．このとき「観光スポット」を中心にした円周上に「店舗」を表示することで，観光スポットから見た店舗の方角を示し，観光スポットから距離が近い店舗ほど大きく，遠い店舗ほど小さく表示することで距離を示す．これにより，複数の「店舗」を Pick し,「観光スポット」へ Lap すると,「観光スポット」を中心とした円周上に複数の「店舗」が一度に表示され，相互の位置関係を一目で確認可能となる．

4. 意味的アフィン変換

コンテンツ融合モデルを形式化するために，コンテンツ・セグメントの融合を幾何的に表現する意味的アフィン変換を提案する．Pick and Lap 操作によるコンテンツの融合は，重ね合わせられる複数のコンテンツ・セグメントに対して，それらの意味的な関係に基づいて，拡大・縮小等の幾何学的変換を行い，適切な配置で，適切な順番で重ね合わせることによって実現する．

2 次元アフィン変換は 2 次元空間上の代表的な幾何変換である．2 次元アフィン変換は，以下のような行列演算として表現できる．

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

ここで，(x, y) は変換前の座標であり，(x', y') は変換後の座標である．

いま，2 次元アフィン変換の変換行列を T とし，重ね合わせの対象となる 2 次元コンテンツを A， B とするとき，Pick and Lap 操作に基づく情報の融合は次式のように形式化できる．

$$Lap(Pick(A), B) = Over(Conv(T, A), B) \tag{2}$$

ここで，$Pick(A)$ は A を Pick することによって得られるコンテンツ・セグメントを表し，$Lap(C, B)$ はコンテンツ C を B 上に Lap することを表す． $Over$ は画像を重畳させる関数であり $Conv(T, A)$ はコンテンツ A を 2 次元アフィン変換行列 T に基づいて幾何変換を行うことを表す．

意味的アフィン変換では，対象となるコンテンツの意味的な関係に基づいて T を決定する．なお，簡単化のために，幾何変換として，縦横比を維持した拡大・縮小と平行移動のみを考える．このとき，2 次元アフィン変換は以下の式で表現できる．ここで，t は拡大・縮小率であり，α は X 方向の移動距離，β は Y 方向の移動距離を表す．

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} t & 0 & \alpha \\ 0 & t & \beta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

重ね合わせられるコンテンツ・セグメントの組み合わせによって変換行列 T は変化する．前章ではユーザの意図を考慮した提示パターンについて述べたが，それぞれの提示パターンの表現を行うために，式 (3) の変換行列 T を求める式を示す．

© 2010 Information Processing Society of Japan

4.1 コンテンツ・セグメントに含まれる画像の種類

ここではコンテンツ・セグメントに含まれる画像を以下の 2 種類に分けて考える．

(1) 意味的な広がりを持つ画像

(2) 意味的な広がりを持たない画像

意味的な広がりを持つ代表的な画像として地図がある．地図は，どこの場所の地図であるかを示す緯度・経度，縮尺等の意味的なデータを保持し，また，平面的な広がりの中で位置ごとに地域の場所を示すように，緯度・経度を軸とした座標系を持っており，意味空間を表す画像といえる．地図以外の例として時刻表，カレンダーなどがあげられる．

意味的な広がりを持たない画像の代表的な例としては，店舗画像がある．店舗画像は店舗の位置情報や店情報など意味的なデータを保持しているが，地図とは異なり，空間的な広がりの中で位置ごとに意味的な情報が対応していないと考えられるため，アイコンとして捉えることができる．

コンテンツ・セグメントに含まれる画像の種類の違いによって，重ね合わせた際に異なるコンテンツ融合モデルを適用する．例えば，店舗画像を地図画像へ重ね合わせた場合には，店舗画像をアイコンとしてとらえ，地図画像が有する意味的な空間の中で，店舗画像を縮小して，店舗が存在する位置に表示する．一方，店舗画像のように意味空間を持たない画像同士を重ね合わせた場合は，相互の関連性を表すことのできる軸を与え，新しく仮想的な意味空間を構成する必要がある．

4.2 意味的な広がりを持つ画像と持たない画像の組み合わせ

4.2.1 店舗を Pick して地図へ Lap

重ね合わせの対象となるコンテンツとして，店舗紹介と地図を考える．図 5 のように，店舗のメタデータは $(lat_s, lng_s, width_s, height_s, img_s)$ であり，地図のメタデータ $(lat_{min}, lng_{min}, lat_{max}, lng_{max}, width_m, height_m, img_m)$ であるとする．ここで，lat_s は店舗の緯度，lng_s は店舗の経度を表し，$width_s$ と $height_s$ は店舗画像の幅と高さを表し，img_s は店舗画像の画像を表す．また，lat_{max} と lat_{min} は地図に表現された緯度の最大値と最小値を表し，lng_{max} と lng_{min} は地図が表す緯度の最大値と最小値を表し，$width_m$ と $hight_m$ は地図画像の幅と高さを表し，img_m は地図画像を表す．

「店舗」を Pick して「地図」へ Lap した場合，ディスプレイ上には図 6 のようにカメラを通して見た地図上の適切な位置に店舗が表示される．つまり，Pick した店舗画像を Lap した対象の地図上の適切な位置 (α, β) に表示するとしたとき，この (α, β) は，店舗の緯度・経度 (lat_s, lng_s) を地図画像の左下端を基準にして，携帯端末のディスプレイ上での地図画像の矩形の幅 $(width_d, height_d)$ を利用して，以下の式により求める．

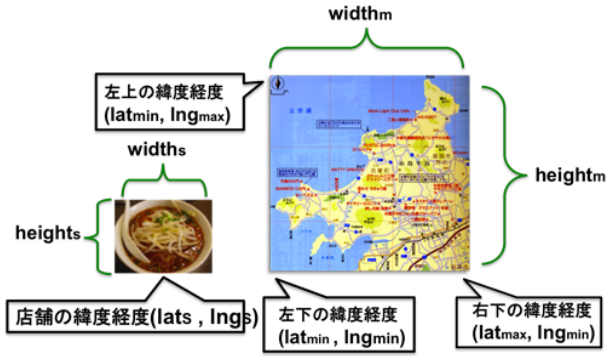

図 5 コンテンツ・セグメントとメタデータ

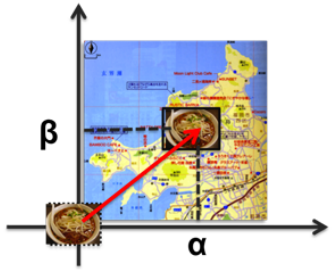

図 6 店舗を地図へ重ね合わせた場合

図 7 地図を店舗紹介へ重ね合わせた場合

$$\alpha = \frac{lat_s - lat_{min}}{lat_{max} - lat_{min}} \times width_d \quad \beta = \frac{lng_s - lng_{min}}{lng_{max} - lng_{min}} \times height_d \tag{4}$$

また，店舗画像の大きさは，携帯電話の画面上での地図の大きさによって変化する．携帯電話の画面上での地図の大きさに対する店舗画像の大きさの割合を s としたとき，店舗画像の拡大縮小率は $t = (s \times width_d)/width_s$ と表せられる．以上の α，β，t は式 (3) のように変換行列のパラメータとして表すことができる．

© 2010 Information Processing Society of Japan

Vol.2010-DBS-151 No.26
2010/11/13

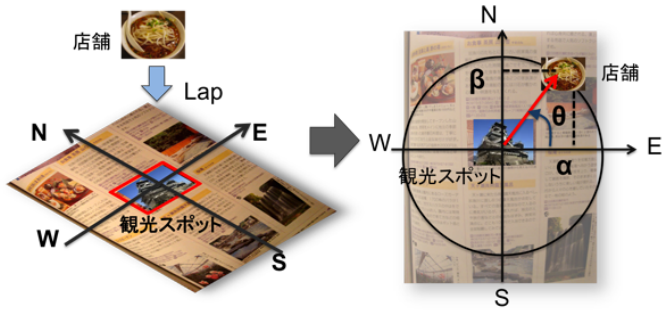

図 8 店舗を観光スポットへ重ね合わせた場合

4.2.2 地図を Pick して店舗へ Lap

「地図」を Pick し「店舗」へ Lap すると,「店舗」側の紙面上では,「店舗」の画像領域の中心を店舗の緯度経度に一致させる． すなわち，図 7 のように,「地図」を,「店舗」の画像領域の中心が地図上における店舗の位置と一致するように，(α, β) 移動させる．このとき，α と β を求める式は

$$\alpha = -\frac{lat_s - lat_{min}}{lat_{max} - lat_{min}} \times width_d \quad \beta = -\frac{lng_s - lng_{min}}{lng_{max} - lng_{min}} \times height_d \tag{5}$$

と表すことができる．

4.3 意味的な広がりを持たない画像を持つコンテンツ・セグメントの重ね合わせ

「店舗」を「観光スポット」へ重ね合わせた場合のように，意味的な広がりを持たない画像を持つコンテンツ・セグメント同士を重ね合わせた場合について述べる．

アイコンとして捉えられる画像を含むコンテンツ・セグメント同士を重ね合わせる場合は，コンテンツ・セグメント間の関係に基づいて，Lap されるコンテンツ・セグメント上で意味的空間を決定する．つまり,「店舗」を「観光スポット」へ重ね合わせた場合，相互の位置関係を表すために,「観光スポット」に含まれる画像を中心にして方角を示す軸を仮想的に作ることで,「観光スポット」を含んだ紙面上に方角を示す意味的空間を与える (図 8). このとき,「観光スポット」を中心にした円周上に「店舗」を表示することで，観光スポットから見た店舗の方角を示すための式を以下に表す． 観光スポットから店舗までの緯度経度の変位をそれぞれ $\alpha' = lat_{shop} - lat_{spot}$,$\beta' = lng_{shop} - lng_{spot}$ とする．ここで，lat_{shop} と lng_{shop} は店舗が存在する緯度と経度を表し，lat_{spot} と lng_{spot} は「観光スポット」の，半径 R の円周上に「店舗」を表示するとしたとき， 「店舗」は「観光スポット」の矩形画像を中心として（α,β）移動し， それぞれは

$$\alpha = Rcos\theta = R \times \frac{\alpha'}{\sqrt{\alpha'^2 + \beta'^2}}$$

$$\beta = Rsin\theta = R \times \frac{\beta'}{\sqrt{\alpha'^2 + \beta'^2}} \tag{6}$$

と表される． ここで円周の半径 R はカメラからの距離によって決まり， カメラを遠ざけるほど R は大きくなる．また相互間の距離を画像の大きさで表し， 距離が離れているほど小さく表示する．この画像の縮小率は

$$t = 1 - \frac{r}{A} = 1 - \frac{\sqrt{\alpha'^2 + \beta'^2}}{A} \tag{7}$$

と表される． A は観光スポットからの任意の最大距離とする．

5. プロトタイプシステム

提案手法のプロトタイプシステムについて述べる．現在，提案手法の有効性を検証するためのプロトタイプシステムを開発している．本プロトタイプシステムは，パーソナルコンピュータ上で USB カメラを利用して動作する (図 9). 店舗や地図の矩形画像は、あらかじめ取得して、関連するメタデータと共に登録しておく．Pick，Lap 時に、撮影領域内の矩形内の画像が抽出され，登録された画像との類似検索を行う．類似検索には画素値のヒストグラムを利用する．

6. 議　　論

意味的アフィン変換を利用することによって，本論文で述べた以外の重ね合わせのパターンや，他の実空間コンテンツを対象にした場合でも，Pick and Lap 手法を実現することができる．例えば，対象とするコンテンツ・セグメントが「店舗」だった場合,「店舗」が持つ情報には位置情報以外にも，平均価格や開店時間などが含まれる場合がある．こうした「店舗」同士を重ね合わせると，互いの平均価格を比べることができる形で提示することも

© 2010 Information Processing Society of Japan

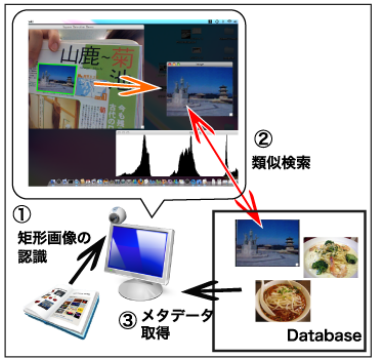

図 9 プロトタイプシステム

可能である．このとき Lap される「店舗」を中心に平均価格を表す仮想的な軸を作り，図のように Pick された「店舗」の矩形画像を軸上で移動させることで表すことができる（図 10 ）．つまり平均価格を意味的アフィン変換の変換行列のパラメータとして利用することが可能でなる．

また，重ね合わせる対象として，意味的な広がりを持った実空間コンテンツは地図以外にも，コンテンツの空間的広がりが時間的な広がりを表すカレンダーや，時間の流れを表す時計などがあげられる．例えば，カレンダーの場合，一般的に時間や日を枠やマスなどによって空間を区切ることで表されているが，それぞれの位置に応じて対応する日や時間が異なり，カレンダー上での空間的な位置が意味を持っている．人はイベントの日程などをカレンダー上で確認する際には，こうしたカレンダーが持つ意味的な空間を理解しながら情報の融合を行っている．

このように本論文で対象とした実空間コンテンツ以外にも，例えば，道端のポスターを Pick しカレンダーに Lap するとカレンダー上に映画の上映期間が表示されたり，店舗を Pick し時計に Lap すると，店舗の開店時間を時計の円周上に表示させることが可能となると考えられる (図 11).

7. 関 連 研 究

近年，セカイカメラ[5] や実空間透視ケータイ[6] のように，カメラ付き携帯電話を用いた拡張現実に関する研究や開発が盛んに行われいる．その中でも実空間上のコンテンツを拡張

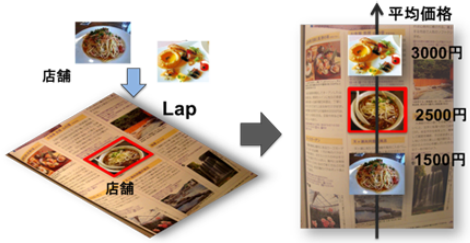

図 10 店舗の平均価格を意味的なデータとした場合

図 11 カレンダーや時計への重ね合わせの例

することに着目した情報閲覧の研究も多く行われている．“HOTPAPER”[7] では，携帯電話を用いた紙媒体への，情報付加，取得を実現している．HOTPAPER では，携帯電話のカメラで紙面を覗いた場所を特定するために，従来のような二次元バーコードなど特別なマーカーを利用せず，文字列の空間的分布の特徴ベクトルを算出することによって，対象とする実空間コンテンツを認識としている．ここでは，使用する紙媒体を任意の大きさの矩形区切りであらかじめ画像として登録しておき，カメラで覗いた際にマッチングで得られた画像に対して情報を付加，または，登録された情報を取得する．“HOTPAPER”は，不動産案内誌や会議のレジュメ，学会のプログラムなど，事前に用意された紙媒体をユーザが印刷して利用することを前提としている．ユーザは会場にて，印刷した紙媒体に対して，携帯電話で現場の写真や，音声，手書きメモなどを仮想的に付加することができる．

Rohs[8] らの研究では，実空間上の地図にあらかじめ，ドットを格子状に配置し，各格子内の 32*32 ピクセルの領域を全て画像として登録しておくことによって，携帯電話のカメラから地図を覗いた際に，画像のマッチングを行い，地図上のどの部分を覗いているか認識可能としている．これを応用した研究に，“Wikeye”[9] や“Real World”[10] がある．“Wikeye”は，携帯電話のカメラを通して，地図上の任意のランドマークを覗くと，ランドマークの情報を記した“WikiPedia”のページが，ディスプレイ上に表示される．さらに“Real World”では，同様にドットをあらかじめペーパークラフトに記しておくことによって，地球型の 3 D オブジェクトを携帯電話のカメラを通して覗くと，覗いた場所の地名や山脈の情報を記した“WikiPedia”のページが表示される．

“CaMus”[11] の研究では，携帯電話とマーカーシートの位置関係から，携帯電話の移動や，回転，高さなどの動きの情報を取得し，その情報を MIDI データへ変換することによって，インタラクティブな音楽演奏を実現している．マーカーシートは，6*6 ドットから構成

© 2010 Information Processing Society of Japan

されるセルがシート全体に配置されている．各セル内には，白と黒のドットが規則的に記されており，これらのパターンの違いを携帯電話のカメラから認識して，携帯電話とマーカーシートとの位置関係を認識している．

これらの研究では，任意の実空間上のコンテンツに付加情報を対応づけ，それらの情報を携帯電話のカメラから取得している．このため情報を取得する閲覧者にとっては，あらかじめ対応づけられた，コンテンツと付加情報の関係のみの閲覧に限られている．本研究では，実空間上のコンテンツを他の実空間上のコンテンツへ重ね合わせることで，閲覧者の選択の違いによって，取得される情報が変化する手法を提案している．また上記の研究はいずれも，携帯電話を，実空間上のコンテンツに対する Peephole(覗き穴) ディスプレイ[12] として機能させており，ユーザにとって直感的な操作で閲覧することを可能 としている．本研究では，任意のコンテンツを他のコンテンツへ重ねた (Lap) 際に，Peephole と同様のアプローチで，携帯電話のディスプレイ上に情報を提示する．

一方で，ユーザの行動から意図に合った情報を推薦する研究も盛んに行われている．小寺ら[13] の研究では、Yahoo Maps[14] のようなオンライン地図を対象に，ユーザの地図に対する領域操作と，地図上の地理オブジェクトに対してカテゴリ検索を行う際のカテゴリ選択の推移からユーザの意図を抽出し、地域情報の重要度を算出することで，ユーザに対して効果的な地域情報を提示している．また実世界のオブジェクトに対するユーザの行動に着目した研究に前川ら[15] の研究がある．ここではユーザがオブジェクトを使った際に，オブジェクトに添付されたセンサノードを通して，そのオブジェクトを対象にした行動に関する知識やノウハウを含む Web ページを取得することを実現している．

本研究では，ユーザが日常の活動の中で，意味的に関連する複数の実空間上のコンテンツを相互に参照しながら，必要な情報を取得している点に着目し，それらを，カメラ付き携帯電話を用いて重ね合わせるというユーザの操作によって，コンテンツが保持する意味的なデータに基づいて，コンテンツを融合させる仕組みを提案した．

8. ま と め

本研究ではカメラ付き携帯電話を用いることを前提に， 実空間上で複数のコンテンツを重ね合わせた際に， 相互の意味的な関係を可視化する Pick and Lap 手法を提案した． さらに重ね合わせの組み合わせによるユーザの意図に合った提示方法を示し， 情報融合の一般化へ向けて従来のアフィン変換に意味的データを取り入れたモデルの構築を行った． 今後は本システムのプロトタイプシステムを完成させ有効性を評価する． 今後の課題としては，実空間コンテンツの組み合わせの多様性を考慮し，融合モデルのバリエーションを増やすことによって，他の様々な実空間コンテンツへの拡張を行っていく必要性がある．

参 考 文 献

1) GMAP，available at http://gmap.jp/gmap/
2) ALPSLAB，available at http://www.alpslab.jp/
3) ゼンリン いつも NAVI，available at http://www.zenrin-datacom.net/mobile/
4) R.T.Azuma, ”A Survey of Augmented Reality”. Presence,Vol.6 No.4,pp.355-385, 1997.
5) ”セカイカメラ”, available at http://sekaicamera.com/
6) ”実空間透視ケータイ”, available at http://kazasu.mobi/
7) B. Erol, J. Graham, E. Antunez, J. J. Hull, “ HOTPAPER demonstration: multimedia interaction with paper using mobile phones ”. (M. M2008) pp. 983-984, 2008.
8) M.Rohs, J.Schoning, A. Kruger, and B.Hecht, ”Towards real-time markerless tracking of magic lenses on paper maps”. In Adjunct Proceedings of the 5th Intl.Conference on Pervasive Computing (Pervasive), pp.69-72,2007.
9) B.Hecht, M.Rohs, J.Schoning, A.Kruger, ”Wikeye - using magic lenses to explore georeferenced Wikipedia content”.In Proceedings of the 3rd International Workshop on PER- MID2007.
10) J.Schoning, M.Rohs, A.Kruger,“ Mobile Interaction with the ‘Real World’”.(MobileHCI2009) no.106,2009.
11) M.Rohs, G.Essl,“ CaMus2 - Collaborative Music Perfor- mance with Mobile Camera Phones ”.(ACE2007) pp.190- 195,2007.
12) G.W.Fitzmaurice,“ Situated information spaces and spa- tially aware palmtop computers ”.(Communications of the ACM) Vol.36, No.7, pp.39-49,1993.
13) 小寺, 北山, 角谷,“オンライン地図におけるユーザの地図操作とカテゴリ選択に基づく地域情報提示方式”(DEIM Forum 2009) pp.C2-5,2009.
14) ”Yahoo! Maps”, available at http://map.yahoo.co.jp/
15) 前川, 柳沢, 岸野, 亀井, 櫻井,“実世界行動のための Web 検索手法の提案”. (DEWS2008) pp. B1-1, 2008.

© 2010 Information Processing Society of Japan