

画像オブジェクト間関係に基づく 3次元構造の推定

石井徹・青木恭太

宇都宮大学大学院工学研究科

〒321-8585 栃木県宇都宮市陽東7-1-2

TEL : 028-689-6247 FAX : 028-689-6248

E-mail : ishii@aok.is.utsunomiya-u.ac.jp kyota@is.utsunomiya-u.ac.jp

一枚の画像からの3次元構造推定のために、画像から抽出される境界と領域の両者を用いる。我々は、画像において視点や3次元オブジェクトの位置や向きなどの変化に対して保存されやすい画像オブジェクト間関係を3次元構造推定に用いる。画像オブジェクトの間に形成される個々の画像オブジェクト間関係のみでは、画像オブジェクトの周囲の状況を十分に記述できない。そこで、3画像オブジェクトを要素として、基本関係の重複組み合わせである複合関係を用いる。3次元シーン全体ではなく部分の様相からその部分構造の3次元解釈を与える。接続関係に着目した隣接オブジェクトグループにおける複合関係分布に基づく3次元構造推定法の有効性を検討する。

キーワード : 3次元構造、隣接オブジェクトグループ、複合関係、線分オブジェクト

Estimation of 3-D Structure using the Relations among Image Objects

Toru Ishii・Kyota AOKI

Graduate School of Engineering, Utsunomiya University

7-1-2 Youtou, Utsunomiya, 321-8585 Japan

TEL : 81+28-689-6247 FAX : 81+28-689-6248

E-mail : ishii@aok.is.utsunomiya-u.ac.jp kyota@is.utsunomiya-u.ac.jp

We have used both of edges and regions to understand 3-D structure from a 2-D image. We use the relations between image objects, edges or regions to estimate 3-D structure. If we only use each relations between objects, we can't describe a feature around the object, enough. We form a combination of the relations among three objects. We can understand informations about a combination of the relations. We define the relation about an object to against various troubles in the natural images. We examine the method for estimation of 3-D structure, 'adjacent group'.

keyword : 3-D structure, adjacent group, complex relation, line segment

1 まえがき

自然画像あるいは模擬自然画像から3次元シーンを解釈する。画像を構成する要素として、境界と領域がある。それぞれが互いに画像上の異なる様相を表現している。よって、境界や、領域のどちらか一方の情報を利用するよりも境界と領域の両方を利用することが望ましい。本報告では境界と領域を、境界または領域と領域の3次元構造を推定するために用いる。野村の3次元物体認識の研究においても、境界と領域の両方を用いることにより成果を挙げている[1]。自然画像からノイズやその他様々な要因により、シーン全体において矛盾なく3次元オブジェクトの境界や3次元面を代表する領域を抽出することは難しい。また3次元シーンに対応する画像においては、視点や3次元オブジェクトの向きや位置により画像オブジェクトが得られる状況は様々であるため、画像オブジェクト自身の特徴量はそれに伴い大きく変化する。そのため、境界、あるいは領域と領域の3次元構造を、視点やオブジェクトの向きや位置の変化にも保存されやすい画像オブジェクト間の関係の分布に基づいて推定する。shape from shadingの手法[2]とは推定の基にしているものが異なるため、併用による、認識率の向上が期待できる。

我々は、1画像で画像上での局所領域の特性から画像オブジェクトに対し3次元解釈を与える。ステレオ画像解析では推定に用いる2枚の画像間の対応点決定により2画像間のズレを得て、距離を知ることができる。それをもとに、片方の画像でしか見えていない部分の3次元構造を推定することができる。

我々は単一の画像を使用して、画像を構成する要素から得られる境界と領域の両方を用いて3次元シーンを解釈する。また、画像より境界や領域の得られる状況は様々であるため、視点やオブジェクトの向きといった各種変化に対して保存されやすい画像オブジェクト間の関係の分布に基づいて境界、あるいは領域と領域の3次元構造の推定を行う。さらに、自然画像中に存在する様々な困難にも関わらず、画像の局所的な様相からその部分構造に3次元

解釈を与える。我々は3次元シーンの部分構造理解に有効な画像オブジェクト間関係を定める。本報告では画像オブジェクト間関係分布を構成するために用いる画像オブジェクトの集合を、基準となる線分あるいは領域と領域に、画像上で直接接続する画像オブジェクトの集合とする。この集合を隣接オブジェクトグループと呼ぶ。我々は、隣接オブジェクトグループにおける複合関係分布に基づく3次元構造推定法の有効性を検討する。

2 画像オブジェクト相互間関係分布に基づくシーン記述

2.1 2オブジェクト間関係分布を用いたシーンの記述

画像を構成する要素として、画像上の境界、あるいは画像上で滑らかに値の変化する画素集合として領域がある。これらは互いに画像上の異なる様相を表現している。境界は一般的な画像理解の手掛かりとして、頻繁に使用される。一方、領域は境界では表現できない画素の集合の様相を表現することができる。よって、境界を代表する線分や、滑らかな画素集合である領域のどちらか一方の情報を利用するよりも、線分と領域の両方を利用した方が良いと考えられる。本報告では線分と領域を、線分または領域と領域の3次元構造を推定するために用いる。また、それぞれを線分オブジェクト、面オブジェクトと呼び、さらに、これらを合わせて基本オブジェクトと呼ぶ。

3次元シーンに対応する画像においては、視点や3次元オブジェクトの向きや位置により画像オブジェクトが得られる状況は様々であるため、画像オブジェクト自体の特徴量はそれに伴い大きく変化する。そこで、視点の変化や3次元物体の向きの変化に大きくは依存しない画像オブジェクト間の関係の分布に基づいて線分、あるいは領域と領域の3次元構造推定を行う[4]。基本オブジェクトは2種類であるから、基本オブジェクト間関係は以下の3つに分類できる。

- 線分オブジェクト相互間の関係
- 面オブジェクト相互間の関係
- 線分オブジェクトと面オブジェクトの関係

線分オブジェクト間の関係は、互いが接続関係にあるか否か平行か否か、さらにそれぞれの端点の位置などを条件として分類する。これには無関係も含まれる。面オブジェクト間の関係は、互いに隣接しているかどうか、互いが包含関係にあるか否かを条件として分類する。これには無関係も含む。線分オブジェクトと面オブジェクト間に存在する関係は、まず線分の面への含まれ方（全て、一部、無）によって大きく3つに分類し、さらに線分の端点の位置、線分と面の境界との接触状況などにより、さらに細かく分類する。

我々は、各画像オブジェクトの性質を関係で記述する、それぞれの画像オブジェクトを含む分布として扱う。各々の基本オブジェクト相互間関係分布のみによるシーンの見え方記述に基づいて線分、あるいは領域と領域の3次元構造推定を行う。しかし、基本オブジェクト相互間関係分布のみによるシーンの見え方記述では、注目する線分あるいは領域と領域がシーン中において対応する部分の3次元構造を特徴づける情報を得ることは困難である。なぜなら、基本オブジェクト間関係は2つのオブジェクトの間に形成されるものであり、これは画像オブジェクトの周囲の状況を十分に表現することが出来ないためである。

2.2 複数オブジェクト間の関係組合せを用いたシーンの記述

基本関係のみを用いた分布では、基準とした画像オブジェクトの周囲の状況を十分に記述することができず、その3次元構造を明らかにすることは難しい。そこで、3つ以上のオブジェクトを要素として、基本関係の組合せである複合関係を形成すると、同一の基本関係を仲立ちとして任意の数のオブジェクトについて関係組合せに関する情報を得ることができ、より注目画像オブジェクトの特徴をつかむこと

ができる。我々は、複合関係を用いて3次元シーンの記述を行う。オブジェクト間関係を形成するために、その要素となる基本オブジェクトの数を n とする。オブジェクト間の関係は、2つの基本オブジェクトの間に形成されるため、 $n > 2$ である。以下に複合関係の定義を示す[4]。

定義1 複合関係

オブジェクト数 n の複合関係は、 n 個の基本オブジェクトのうち、任意の2個のオブジェクトの間に存在する基本関係の重複組合せである。

一つの複合関係を形成するとき、関係を形成するオブジェクトの数が n 個ならば、そこでつくられる基本関係の数は ${}_nC_2$ である。よって、形成される全ての複合関係を考えると、莫大な数の関係を扱うことになり、我々の持つ計算資源の面から考えて扱いが困難である。そこで、形成される関係数を我々が扱う数に制限する必要がある。よって、複合関係形成のための要素数を3に限定する。

3 画像上の局所領域における画像オブジェクト相互間の関係分布を用いたシーン記述

3.1 画像上の局所領域における2オブジェクト間関係分布を用いたシーンの記述

3次元シーンと対応する画像において、画像オブジェクト固有の特徴量は、視点や3次元オブジェクトの向きや位置に対して保存量ではない。一方、画像オブジェクト間の関係分布は、視点の変化や3次元物体の向きの変化に対して保存量である。そこで我々は画像オブジェクト間関係の分布に基づいて線分、あるいは領域と領域の3次元構造の推定を行う。

析の対象とする3次元シーンの投影としての自然画像は、ノイズなどの様々な困難により、3次元オブジェクトの全体像を捉えることができない場合が

多い。そこで画像全体ではなく、画像上の局所的な範囲において、3次元シーンの構造理解に対して有効な画像オブジェクト間関係を定める。

線分がシーン中において対応する部分の両側に広がる3次元面がどのような面であるか、あるいは領域と領域がシーン中で対応する部分は、どのような3次元面であるかを探るためには、その周囲の状況を観察することが重要である。しかし任意に基本オブジェクトを選択して画像オブジェクト間関係を形成すると、平行関係など、距離の離れたオブジェクト間にも成立する関係も含んでしまう。これら画像オブジェクト間関係の形成を、距離の離れたオブジェクト間に成立する関係も含めて行うことは、3次元シーンの部分構造理解に対して有効ではない。単純に基準となる画像オブジェクトを含む画像オブジェクト間関係を全て用いることはできない。

本報告では画像オブジェクト間関係分布を構成するために用いる画像オブジェクトの集合を、基準となる線分あるいは領域と領域に、画像上で直接接続する画像オブジェクトの集合とする。なぜならば、画像上における画像オブジェクトは大きさが様々であるため、オブジェクトの大きさに基づいたオブジェクト集合の定義は困難であり、オブジェクトレベルでは、接続関係のみが考えられるからである。この集合を隣接オブジェクトグループと呼ぶ。以下に隣接オブジェクトグループの定義を示す[5]。また、隣接オブジェクトグループの例を図1に示す。図中に、矢印で指し示した線分を基準オブジェクトとし、これに注目する。基準オブジェクトは、0隣接オブジェクトでもあるため、これより1隣接オブジェクトの集合を導く。つまり、0隣接オブジェクトと接続関係にある基本オブジェクトは、図中の4つの線分と4つの面である。これら8つの基本オブジェクトが1隣接オブジェクトとなる。また、0隣接オブジェクトと1隣接オブジェクトそれぞれの集合の和をとることで、1隣接オブジェクトグループが形成される。

定義2 隣接オブジェクトグループ

ある基本オブジェクトに注目し、これを基準オ

ブジェクト k とする。さらに、基準オブジェクトを0隣接オブジェクトと呼ぶ。ここで、 $\forall i < n$ とすると、 i 隣接オブジェクトの集合 ${}_kR_i$ が得られているとする。このとき n 隣接オブジェクトの集合 ${}_kR_n$ を $\{x \mid x \text{は} \exists y \in {}_kR_{n-1} \text{と接続関係を持ち、かつ} x \notin {}_kR_j; \forall j < n\}$ とする。また、 n 隣接オブジェクトグループを ${}_kR'_n = \cup_{i=0}^n {}_kR_i$ と定義する。

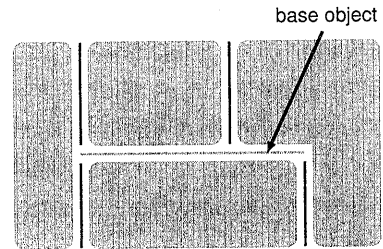


図 1: 隣接オブジェクトグループの例

我々は、隣接オブジェクトグループにおける2オブジェクト間の関係分布を用いてシーンの記述を行う。隣接オブジェクトグループにおける基準オブジェクトに関わる2オブジェクト間の関係の分布を用いることで、ノイズ等様々な困難を持った自然画像への我々の手法の適用を可能にする。これは、基準画像オブジェクトの周辺との関わりを明らかにしつつ、画像上において距離的に離れたオブジェクト間にも成立する関係を関係分布の集合に含めないことで、画像の局所領域に注目することができるためである。しかし、基本オブジェクト間関係は2つのオブジェクトの間に形成されるものであり、これだけでは基準オブジェクトの周囲の状況を十分に表現できないと考えられる。

3.2 画像上の局所領域における複数オブジェクト間の関係組み合わせを用いたシーンの記述

隣接オブジェクトグループにおける基本関係のみを用いた分布では、画像上における局所領域への注目を可能にするものの、基準オブジェクトの周囲の状況を十分に記述できず、その3次元構造を十分に

記述できない。そこで、隣接オブジェクトグループにおける要素数3の複合関係を形成することで、同一の基本関係を仲立ちとした関係組み合わせに関する情報を得ることができ、より基準オブジェクトの特徴を掴むことができると考えられる。そこで我々は、隣接オブジェクトグループにおける複合関係分布を用いてシーンの記述を行う。

4 線分の3次元構造推定実験

4.1 学習画像を用いた線分の3次元構造推定

我々は、隣接オブジェクトグループにおける複合関係分布による、画像オブジェクトの3次元構造推定法の有用性を示すために、オブジェクトを対象として推定実験を行った。実験は次のものを行う。注目線分オブジェクトを含む基本関係の分布による線分の3次元構造推定、注目線分オブジェクトを含む複合関係の分布による線分の3次元構造推定、および隣接オブジェクトグループにおける注目線分オブジェクトを含む基本関係の分布による線分の3次元構造推定と提案手法、の4つである。

文字認識などの研究と異なり、本手法で推定の手がかりとする複合関係は種類の数膨大であり、それら全てに対してシーン解釈のためのモデルを入手で作成することは困難である。そこで学習データに基づいた推定を行うことにする。

学習データからは、隣接オブジェクトグループごとの複合関係分布と線分オブジェクトの3次元構造を得る。隣接オブジェクトグループ中に含まれる画像オブジェクト数は、基準となる画像オブジェクトの大きさなどの影響で大きく変動する。この変動の影響を避けるために、隣接オブジェクトグループ中において複合関係を扱うときには、その数ではなく、複合関係の割合を用いることにする。本報告では、学習データに基づいた線分の3次元構造推定を、隣接オブジェクトグループ中の複合関係の割合を用いて行う[5]。推定の方法としては、分布と3次元構造との相関を用いる。

式1に推定値算出のための式を示す。

$$E(o, L) = \sum_{c \in C} \left(\frac{r(c, o) - Av(c)}{Sdv(c)} \right) \times S(c, L) \quad (1)$$

但し、式中の変数は以下に示すものとする。

o	画像オブジェクト
L	3次元構造ラベル
C	すべての複合関係
$r(c, o)$	画像オブジェクト o の隣接オブジェクトグループ中の複合関係 c の割合
$Av(c)$	学習データにおける複合関係 c の割合の平均
$Sdv(c)$	複合関係 c の標準偏差
$S(c, L)$	基準オブジェクトが存在する隣接オブジェクトグループ中の複合関係 c の割合と学習データの基本オブジェクトのラベル L の有無の相関係数
$E(o, L)$	画像オブジェクト o がラベル L を持つ度合いの推定値

式1により、基準オブジェクトに対し、それが各々の3次元構造である可能性が推定値として算出される。線分、あるいは領域と領域に対し、用意した3次元構造の数だけの推定値が算出される。このうち、推定値が最大のものを推定3次元構造とする。

4.2 推定実験

今回行った実験を以下に示す。

実験1 注目線分オブジェクトを含む基本関係の分布による線分の3次元構造推定

実験2 注目線分オブジェクトを含む複合関係の分布による線分の3次元構造推定

実験3 隣接オブジェクトグループにおける 注目線分オブジェクトを含む基本関係の分布による線分の3次元構造推定

実験4 隣接オブジェクトグループにおける注目線分オブジェクトを含む複合関係の分布による線分の3次元構造推定

表1に示す3つの学習画像グループを用意した。線分は3次元の意味を持っていないので、3次元特性を示す情報が必要である。線分が3次元シーンにおいて対応する部分の両側に広がる3次元面の相互関係を分類し、お互いが重なる部分や接触部分の視点からの距離、さらに面相互が平行であるか否かを考慮し、今回以下に示す5つの3次元構造を定義して線分を扱う。また、実験に用いた10枚の画像を図2に示す。

- 1つの3次元オブジェクトに属する、距離の等しい3次元面相互の凸状の接触部分。
- 1つの3次元オブジェクトに属する、距離の等しい3次元面相互の凹状の接触部分。
- 1つの3次元オブジェクトに属する、距離の等しい3次元面相互の滑らかな接触部分。
- 異なる3次元オブジェクトに属する、距離の等しい3次元面相互の接触部分。
- 異なる3次元オブジェクトに属する、距離の異なる3次元面同士の重なり。

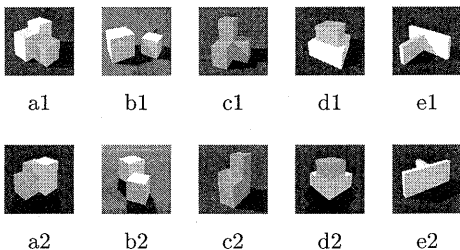


図 2: 実験画像

表 1: 学習画像のグループ分け

グループ	学習画像
group1	a1,b1,c1,d1,e1
group2	a2,b2,c2,d2,e2
group3	a1,b1,c1,d1,e1,a2,b2,c2,d2,e2

実験1. 注目線分オブジェクトを含む基本関係の分布による線分の3次元構造推定

抽出した全ての画像オブジェクトに対し、任意の2オブジェクト間で基本関係を形成し、その分布によって線分の3次元構造を推定する。推定結果を表2に示す。

表 2: 実験1における予測3次元構造の正解率

推定画像 \ 学習画像	group1	group2	group3
a1	0.404	0.128	0.532
b1	0.526	0.132	0.421
c1	0.548	0.262	0.548
d1	0.690	0.143	0.643
e1	0.733	0.111	0.644
合計	0.579	0.154	0.561

group1~3、全てのグループで、正解率の平均が5割を下回った。これは、基本関係では推定する線分周囲の情報を記述することができず、線分の構造特徴が把握できないためであると考えられる。

実験2. 複合関係による線分の3次元構造推定

推定結果を表3に示す。推定する線分の周囲の状

表 3: 実験2における予測3次元構造の正解率

推定画像 \ 学習画像	group1	group2	group3
a1	1.000	0.149	0.979
b1	0.816	0.368	0.684
c1	0.905	0.381	0.881
d1	0.881	0.262	0.810
e1	0.911	0.156	0.844
合計	0.907	0.257	0.846

況を記述できるようになったため、基本関係の場合に比べ、全体的に推定精度が向上している。

実験3. 隣接オブジェクトグループにおける基本関係による線分の3次元構造推定

画像オブジェクトの接続関係に着目した、隣接オブジェクトグループにおける基本関係分布を用いて線分の3次元構造推定を行う。推定結果を表4に示す。

表4: 実験3における予測3次元構造の正解率

推定画像 \ 学習画像	group1	group2	group3
a1	0.638	0.255	0.745
b1	0.579	0.447	0.579
c1	0.548	0.190	0.500
d1	0.690	0.262	0.810
e1	0.711	0.244	0.733
合計	0.636	0.276	0.678

実験4. 隣接オブジェクトグループにおける複合関係による線分の3次元構造推定

画像オブジェクトの接続関係に着目した、隣接オブジェクトグループにおける複合関係分布を用いて線分の3次元構造推定を行う。推定結果を表5に示す。

表5: 実験4における予測3次元構造の正解率

推定画像 \ 学習画像	group1	group2	group3
a1	1.000	0.362	1.000
b1	1.000	0.263	1.000
c1	0.976	0.333	0.952
d1	0.952	0.310	0.976
e1	0.956	0.244	0.956
合計	0.977	0.304	0.977

4.2.1 実験結果からの提案手法の考察

隣接オブジェクトグループを用いない、画像オブジェクト間関係のみを用いた線分の3次元構造推定

と比べ、実験2、実験3、実験4共に推定精度が高くなっている。画像オブジェクト間関係のうち、接続関係に注目したことにより線分の特徴把握に有用でない関係を除くことができたためであると考えられる。

隣接オブジェクトグループにおける複合関係分布を用いることで、線分の3次元構造を推定する。今回行った4つの実験の比較を行う。図3に各実験における予測3次元構造の正解率を示す。

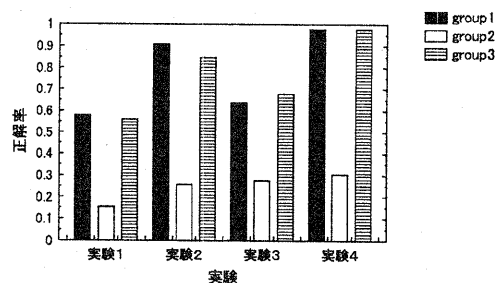


図3: 実験1、2、3、及び4における予測3次元構造の正解率

基本オブジェクト間の関係による推定については、図3の実験1の欄に示したように、基本関係のみによる線分の3次元構造推定は困難であると考えられる。これは、基本関係のみによるシーンのモデルでは、線分の構造を特徴付ける関係分布が得られないためであると考えられる。

注目線分の周囲の状況を記述し、より線分の3次元構造を掴むための手法が必要である。そこで、複合関係による線分の3次元構造推定を行った。前述のとおり、その要素数を3に限定する。図3の実験2の欄に示したとおり、基本関係のみの推定に比べ、正解率が向上していることが分かる。画像オブジェクト間の関係を用いた推定法では、基本関係のみによる方法に比べ、複合関係を用いた方法の方が、より適していることが分かる。

各線分における複合関係分布には、距離が離れたオブジェクト間にも成立する関係も含まれる。これらは線分の特徴把握には不適切だと考えられる。そこで、画像オブジェクト関係のうち、接続関係に注

目した隣接オブジェクトグループによる推定を行った。結果は図3の実験3と実験4の欄に示すとおりである。隣接隣接オブジェクトグループにおける基本関係による線分の3次元構造推定は、基本関係のみによる線分の3次元構造推定に比べ、線分の特徴把握には劣ると考えられる、距離の離れた基本オブジェクト間関係が含まれないので、正解率が向上している。さらに、隣接オブジェクトグループにおける複合関係による線分の3次元構造推定では、複合関係のみによる線分の3次元構造推定に比べ、複合関係を形成する画像オブジェクトの候補を1隣接オブジェクトに制限したため、正解率が向上している。

5 むすび

我々は、隣接オブジェクトグループにおける複合関係分布による、画像オブジェクトの3次元構造推定法の有効性を検証した。線分の3次元構造推定実験では、提案手法の他に、基本関係のみによる推定、複合関係による推定、隣接オブジェクトグループにおける基本関係分布による推定を行った。これらの比較をとおして、画像上の局所的な部分における、基本関係の組みあわせである複合関係を用いた推定手法が、画像オブジェクトの3次元構造推定手法として有効であることが明らかにした。

我々は、1画像で画像上での局所領域の特性から画像オブジェクトに対し3次元解釈を与える。ステレオ画像解析では推定に用いる2枚の画像間の対応点決定により2画像間のズレを得て、距離を知ることができる。それをもとに、片方の画像でしか見えていない部分の3次元構造を推定することができる。学習データを、ステレオ画像解析により得た距離情報を用いて作成することが出来る。

今後の課題としては、以下に示すようなことが考えられる。今回の実験では、推定に3次元構造と複合関係分布の相関を用いた。その理由は単純かつ簡単に3次元シーン解釈のためのモデルが作成できるからである。複合関係の割合と3次元構造との様々な関係を分布で捉え、より詳細な解析を行うことが

考えられる。また、ニューラルネットワークを利用して複合関係の割合から、3次元構造を推定することも考えられる。

参考文献

- [1] 野村由司彦、張滌利、坂井田裕子、藤井省三 “陰影とエッジの情報を融合したモデルベースト3次元物体認識”, 電子情報通信学会論文誌, 5J-D II, pp.834-839(1996)
- [2] Vincenzo Caglioti, Eugenio Castelli, "Recovering cylindric and conic surfaces from contours and reflections", Pattern Recognition Letters 20, pp.367-382, 1999
- [3] M.B.Clowes "On Seeing Things", Artificial Intelligence, 2, pp79-116(1971)
- [4] 菊島公弘、青木恭太 “隣接オブジェクトグループにおける3画像オブジェクト間の接続関係組合せによる線分の3次元構造の推定”, 情報処理学会 コンピュータビジョンとイメージメディア 108-8, 1997.11.27
- [5] Kyota AOKI, Kimihiro KIKUSHIMA, Sou HINATA, Toru ISHII "Estimating Three-dimensional Relations between two Faces based on the Relations among Image Objects", International Symposium on Image, Speech, Signal, Processing, and Robotics (ISSPR'98), I, pp187-190(1998)