

ステレオによるシーンの解釈

杉本和英* 高橋裕信* 富田文明**

* 三洋電機筑波研究所

** 電子技術総合研究所

エッジベースステレオでは、まず、ステレオ画像間でエッジの対応を求め、次に、完全な距離画像を内挿によって求めるのが基本的な方法である。最初の対応問題を解決するために、多くのアルゴリズムが報告されている。しかし、いずれも輪郭エッジと稜エッジ（奥行きと方向の不連続部分）を区別しないために、次の内挿の単純な方法では、偽の面を生成してしまう。本論文では、ステレオ画像間での対応の単位として、領域を右側に見るように方向付けられた境界線セグメントを用い、平面物体の存在するシーンに対して面の解釈を行なう方法について述べる。

まず、輪郭エッジを明らかに区別できるオクルージョン点を検出する。次に、他の未確認エッジを、明らかな輪郭エッジを用いて、輪郭エッジまたは稜エッジに分類する。さらに、稜エッジは、両側の面の法線方向に基づいて、凸、凹、平坦（模様や影による）エッジに分類する。そして、最後に、シーンの空間を、内挿された実平面を用いて、自由空間、占有空間、解釈不能空間に分類する。

Scene Interpretation Based on Boundary Representations of Stereo Images

Kazuhide Sugimoto* Hironobu Takahashi* Fumiaki Tomita**

Tsukuba Research Center
SANYO Electric Co., Ltd.

Computer Vision Section
Electrotechnical Laboratory

2-1 Koyadai, Tsukuba-shi, Ibaraki 305 1-1-4 Umezono, Tsukuba-shi, Ibaraki 305

The usual approach of edge-based stereo is to first identify corresponding edges between the two images, then interpolate to obtain a complete distance or depth map. Many algorithms have been proposed to solve the first correspondence problem. However, occluding and ridge edges (depth and orientation discontinuities) have not been explicitly detected and this has made surface interpolation difficult; simple interpolation creates false surfaces. In this paper, we solve the problem for objects with planar surfaces, using stereo vision which is based on directional boundary segments of regions in a pair of stereo images.

We propose methods to detect occlusions at which occluding edges are explicitly identified, to classify the other unidentified edges into either occluding edges or ridge edges, using the explicitly occluding edges, to classify ridge edges further into convex, concave, or flat (paint or shadow) edges, based on the surface normals of both sides of the edges, and finally to classify the space of the scene into free, occupied, or uninterpretable spaces, using interpolated real surfaces.

1 まえがき

ステレオは、シーンの3次元情報を求める最も一般的な方法として期待される。エッジベースステレオでは、まず、ステレオ画像間で対応するエッジを求め、次に完全な距離または奥行きデータを内挿によって求めるのが基本的な方法である。最初の対応問題を解決するために多くのアルゴリズムが報告されている⁽¹⁾。次の内挿は、ステレオでは面上にテクスチャが存在しなければ面の情報を直接得られないためである⁽²⁾。最近、ステレオは、実験的にはあるが、環境を観察し、その距離情報を獲得する移動ロボットに使用されるようになってきた⁽³⁾⁽⁴⁾。しかし、たとえロボットが障害物のエッジを検出可能だとしても、エッジのどちら側が面であるか、つまり、エッジのどちら側を通過できるのか決定することができない。輪郭エッジと稜エッジ（奥行きと方向の不連続部分）を区別していないので、単純な内挿の方法では、偽の面を生成してしまう。

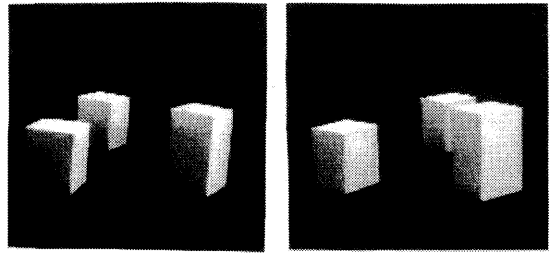
本論文では、ステレオ画像間での対応の単位として、領域を右側に見るように方向付けられた境界線セグメントを用い、平面物体の存在するシーンに対して面の解釈を行なう方法を提案する。このセグメントベースステレオは幾つかの利点を持っている。セグメントは直線または曲線の式により近似されるので、高い精度で3次元セグメントの座標を出力することができる。また、画像中のセグメントの高い連続性とその方向性は、3次元平面や空間の(閉)境界線を形成する3次元セグメントに保存されている。

まず、明らかな輪郭エッジを区別することができるオクルージョンの生じている部分を検出する。次に、他の未確認エッジを、明らかな輪郭エッジを用いて、輪郭エッジまたは稜エッジに分類する。更に、稜エッジは両側の面の法線方向のなす角度によって、凸、凹、平坦（模様や影による）エッジに分類する。そして、最終的にシーンの空間を、内挿された実平面を用いて、自由空間、占有空間、解釈不能空間に分類する。

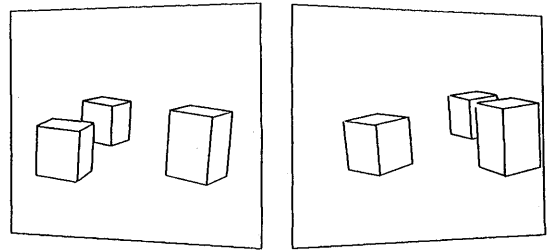
2 データ構造

2.1 画像データ

画像を表現するするために、CADの分野で開発された最も一般的なソリッドモデルである、3次元B-rep (Boundary Representation) の仕様にしたがって、画像の2次元B-repを既に開発した⁽⁵⁾。これは、画像の



(a)



(b)

図1 画像データ：(a)ステレオ画像；(b)境界線セグメント

中間表現として用いられ、もとの画像を再構成するのに十分な情報を持つだけでなく、多くの高レベルの後処理に有効に使うことができる。以下にその概要を述べる。

画像は、エッジ検出法により領域に分割される。各領域は、1つの外周境界線と、幾つかの穴が存在すれば、穴との内周境界線のリストによって表現される。そして、各境界線は、それを構成する境界線セグメントのリストによって表現される。各セグメントは、領域を右側に見るように方向付けられており、始点と終点の座標、および直線の方程式をその値に持つ。図1に、ステレオ画像のB-repから生成した境界線セグメントを示す。ステレオ画像の対応処理は、このデータ構造上で行なわれる。対応の単位として、境界線セグメントが用いられる⁽⁶⁾。

2.2 ステレオデータ

対応する一対のセグメントは、一つの3次元セグメントを表わす。これを、「アーク」と呼ぶことにする。ステレオの出力は、面または空間の(閉)境界線を形成する連結アークのリストである。例として、図2(a)の3次元ブロックを考える。ブロックには輪郭エッジ(→)と凸稜エッジ(+)が観測される。ステレオの出力

は、ブロックを図2(b)の4つの境界線によって表現する。境界線セグメントは実際にはエッジ上に存在するが、各エッジ上には反対向きの2つのセグメントが存在することを示す。各セグメントは、その右側に面が存在することを意味する。しかし、境界線B4上に存在するセグメントは実際には存在しない面を表わす虚セグメントである。輪郭エッジ上には、一対の実セグメントと虚セグメントとが共存する。虚セグメントは、単純な内挿では、偽の面を作る原因となる。そこで、輪郭エッジから虚セグメントを除去することが必要となる。

3 オクルージョンの検出

ステレオにより、両方の画像中に見えている部分の距離を測定することができる。しかし、3次元シーンでは、手前にある物体が後方の物体を隠すことにより、一方の画像中では見えていても、他方の画像中では見えないオクルージョンが少なからずある。オクルージョンの存在が検出できれば、セグメントの不完全な対応の検証ができるのみならず、後述するように、輪郭エッジ上の明らかな「実」セグメントと「虚」セグメントを確認することができる。そこで、まず、ステレオデータからオクルージョンを検出する方法について述べる⁽⁷⁾。

オクルージョンが生じると、図3のセグメント対($S_j, S_{j'}$)のように、対応の不完全なセグメント対が出現する。すなわち、左画像のセグメント S_j のABの部分、右画像のセグメント $S_{j'}$ のA'B'の部分に対応しているが、左画像の領域Rが S_j のBCの部分より手前にあると、右画像では領域Rに対応する領域R'によってBCに対応する部分が隠されて見えなくなる。このようなオクルージョンの検出方法を以下に述べる。

- ① まず、対応の不完全なセグメント対($S_j, S_{j'}$)を探す。
- ② 次に、セグメント $S_{j'}$ の端点 B' を通る他のセグメント $S_{k'}$ を探す。
- ③ S_k は $S_{k'}$ に対応しているとする、点 B' はセグメント S_j 上の点Bと、 $S_{k'}$ 上の点 B'' の2点と対応している。
- ④ Bと B' 間の視差を d_j 、 B'' と B' 間の視差を d_k とする。もし $d_k > d_j$ ならば、セグメント S_j のB

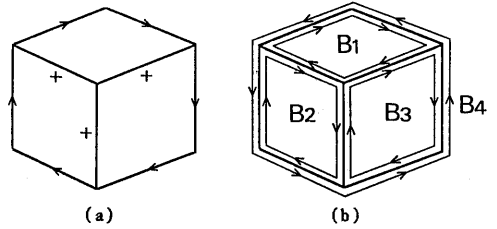


図2 ブロック: (a)エッジ; (b)境界線セグメント.

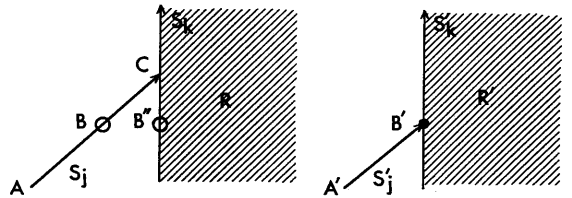


図3 オクルージョン点で不完全に対応するセグメント.

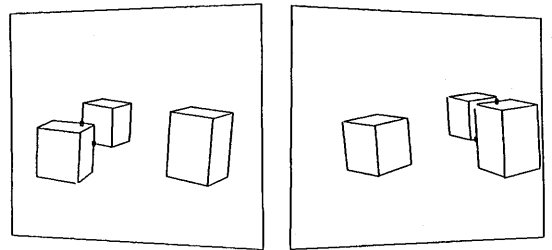


図4 オクルージョン点(O)の検出結果.

Cに対応する部分は、右画像では B' からセグメント $S_{k'}$ (または領域 R')によって隠されていると判定する。

図4は、図1のシーンで検出されたオクルージョン点を示している。

4 セグメントの分類

4.1 実在テスト

オクルージョン点では、図5(a)のように、2つの明らかな「実」セグメントと「虚」セグメントを見つけることができる。そこで、図5(b)のように、「実」セグメントまたは「虚」セグメントを持った境界線上の他の未確認セグメントを、実セグメントまたは虚セグメントに分類する一般手順(実在テスト)を以下に示す。

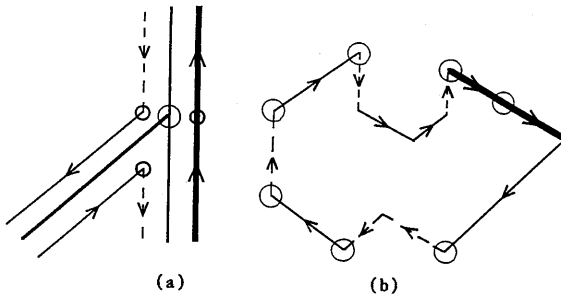


図5 実在テスト: (a)オクルージョン点(O)における「実」セグメント(太線)と「虚」セグメント(点線); (b)境界線上の未確認セグメント(細線)。

- ①すべての「実」セグメントを含み、「虚」セグメントを含まないセグメントの組で平面を構成する。(構成できない場合は、その境界線は解釈不能である。)
- ②境界線上のすべてのセグメントが、平面上または平面上空に存在するような平面が唯一存在すれば、その平面が、実平面であると考え(存在しない場合または複数存在する場合は、その境界線は解釈不能である。)
- ③実平面上のセグメントは実セグメントであるが、実平面上空のセグメントは虚セグメントであり、除去される。

以下の実在テストの適用例では、簡単のために、解釈不能な境界線は存在しないものと仮定する。

[簡例1] ブロック上に他のブロックが存在するシーン(図6)を考える。7つの境界線と、2つのオクルージョンが存在する。

- ①「虚」セグメントが存在しない右回りの境界B1, B2, B3, B5, B6を考える。各境界線上のすべてのセグメントは、平面を構成するならば実セグメントである。
- ②「虚」セグメントが存在する右回りの境界線B4を考える。上のブロックが下のブロック上に存在する場合は、すべての未確認セグメントは、平面を構成するならば実セグメントである(図7(a))。しかし、上のブロックが、下のブロックの上空に存在する場合は、上のブロックの下辺の2つのセグメントは実平面上空に存在することになるので虚セグメントであり、除去される(図7(b))。

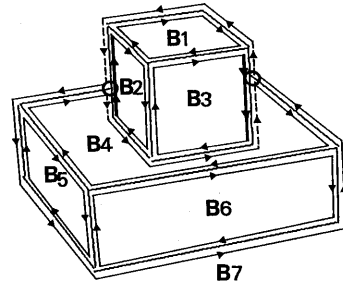
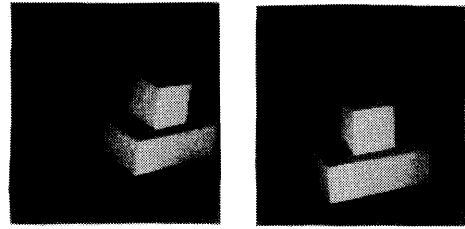


図6 重なったブロック。

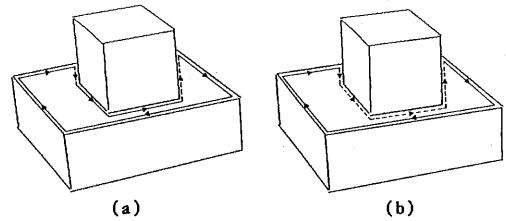


図7 実在テスト: (a)上のブロックが下のブロックの上に存在する場合; (b)上のブロックが下のブロックの上空に存在する場合。

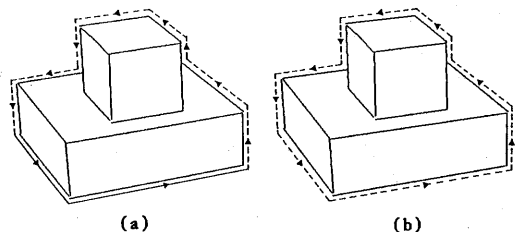


図8 実在テスト: (a)ブロックが床の上に存在する場合; (b)ブロックが空中に浮いている場合。

③「実」セグメントの存在しない左回りの境界B7を考える。右側の周囲の面が床または道路のようにすでに平面として定義されており、下のブロックがその上に存在する場合は、下のブロックの下辺の2つのセグメントは実セグメントである(図8(a))。一方、周囲の空間が自由空間である場合、すなわち、ブロックが空中に浮いている場合は、境界線上のすべてのセグメントは虚セグメントであり、除去される(図8(b))。

【範例2】穴のあいたブロックのシーン(図9)を考える。7つの境界線と、1つのオクルージョンが存在する。

①「実」セグメントが存在する左回りの境界B4は、穴に対する境界線である。言いかえれば、オクルージョン点が存在しない左回りの境界線は、穴とは考えない。この例では、「虚」セグメントは存在しないので、穴の境界上のセグメントは、すべて実セグメントである。

②「虚」セグメントが存在する穴の内側の右回りの境界B5, B6を考える。各境界線は、それぞれ奥の2つのセグメントが実平面を定める(図10)。

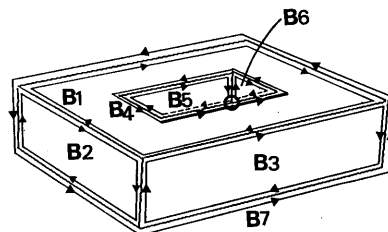
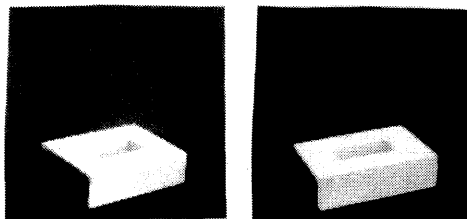


図9 穴の開いたブロック。

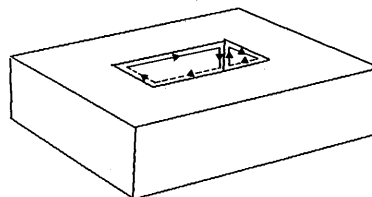


図10 穴の内側。

4.2 解釈不能境界

実在テストは、境界線に対して平面を一意的に定義できない場合には失敗する。そのような境界線は、現在の視点からは解釈不能である。原因としては、多くのオクルージョンが生じた場合や、照明の悪条件などによる、画像でのエッジ検出の失敗である。例えば、図11(a)のように、家の屋根の尾根が見えない場合、たとえ一本のエッジが失われていると仮定しても、図11(b)-(d)のように、その領域について多くの解釈が可能である。このような場合、視点と視線方向を変えて、その原因を知ること、例えば、見えないエッジを検出することが必要となる。

また、两眼視では、水平セグメントの端点の少なくとも一方が隠されている場合、その水平セグメントの距離を求めることができないため、そのセグメントとのオクルージョンを検出することができない問題がある。そこで、水平セグメントについては、その距離の上限あるいは下限を考慮するだけで、直接的には処理の対象とはしないdon't careセグメントとして扱う。

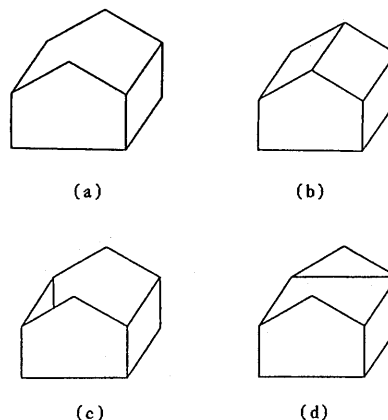


図11 解釈不能な境界線：(a)不完全な家；(b)-(d)可能な解釈。

但し、この問題は、3眼や多眼ステレオによって容易に解決することができる^(*)。

5 エッジの種類

虚セグメントを除去した後、各エッジを、そのエッジ上のセグメントの個数により、輪郭エッジと稜エッジに分類することができる。すなわち、輪郭エッジは1つのセグメントしか持たないのに対して、稜エッジは

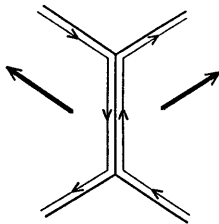


図12 稜エッジの両側の実平面の法線。

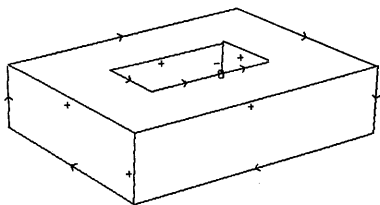
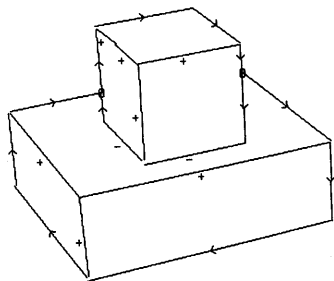


図13 エッジのラベル付けの結果。

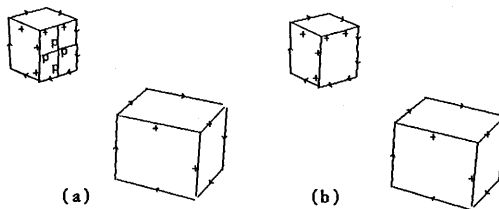
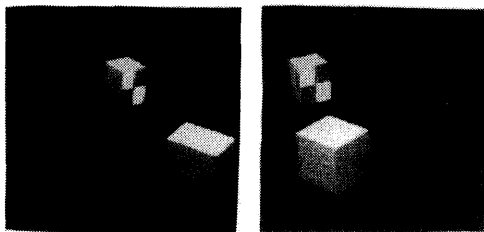


図14 模様エッジ: (a)エッジのラベル付けの結果; (b)模様エッジを除去した後のラベル。

互いに反対向きの2つのセグメントを持つ。

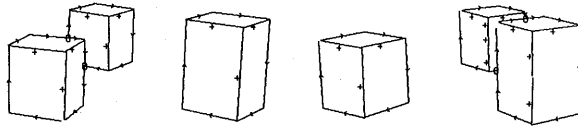
稜エッジはいわゆる翼エッジ(winged edge)であることから、図12のように、その両側の実平面の法線方向のなす角度が、正、負、ゼロによって、凸、凹、平坦(模様もしくは影による)エッジにそれぞれ分類することができる。

以上の処理の結果、各エッジは、輪郭エッジ(→)、凸稜エッジ(+)、凹稜エッジ(-)、平坦稜エッジ(P)に分類される。したがって、図6および図9の各エッジに、Waltzのラベルのように^(*)、図13のラベルを付けることができる。また、平坦稜エッジを除去することによって、図14のように、物体の純粋の立体形状を得ることもできるようになる。

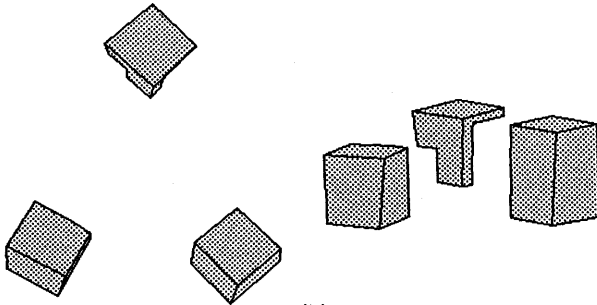
6 空間の種類

最後に、内挿された実平面を用いて、シーンの空間を、それぞれ、自由空間、占有空間、解釈不能空間に分類することができる。図1のシーンを分類した結果を図15に示す。ここで、稜エッジは隣接する2平面の交線により表現されている。

その結果、ロボットは自由空間を移動しながら、占有空間のモデルを更新し、解釈不能空間を解釈可能にすることができるようになる。



(a)



(b)

図15 空間の分類: (a)エッジのラベル付けの結果;
(b)2つの視点から見た空間。

7 まとめ

ステレオ画像の境界線表現に基づいて、3次元シーンの面を解釈する方法を提案した。最初に、明らかな「実」セグメントと「虚」セグメントを区別できるオクルージョン点を検出する。次に、各境界線上の他の未確認セグメントを、実在テストによって実セグメントまたは虚セグメントに分類する。そこで、実セグメントに唯一の平面があてはめられ、その上空の虚セグメントは除去される。その結果、輪郭エッジは1つのセグメントしか持たないのに対して、稜エッジは互いに反対の向きを持った2つのセグメントを持つことから、エッジ上のセグメントの数により、各エッジを、輪郭エッジもしくは稜エッジに分類する。さらに、稜エッジは、その両側の実平面の法線方向のなす角度が、正、負、ゼロによって、凸、凹、平坦エッジにそれぞれ分類する。最後に、内挿された実平面を用いて、空間を、自由空間、占有空間、解釈不能空間に分類する。

本論文では、平面物体に対する方法について述べたが、曲面物体に対しても拡張することができる。今後は、このステレオを備えた移動ロボットによる複雑な環境の世界モデルの構築を予定している。

謝 辞

日頃有益な助言と討論をいただき、電子技術総合研究所視覚システム研究室大島正毅室長はじめ研究室の皆様にご感謝いたします。

参考文献

- ①Barnard, S.T. and Fischler, M.A.: Computational Stereo, Computing Surveys, 14, pp.553-572, 1982.
- ②Hoff, W. and Ahuja, N.: Extracting Surfaces from Stereo Images: An Integrated Approach, Proc. 1st ICCV, pp.284-294, 1987.
- ③Ayache, N. and Faugeras, O.D.: Building a Consistent 3D Representation of a Mobile Robot Environment by Combining Multiple Stereo Views, Proc. 10th IJCAI, pp.808-810, 1987.
- ④Tsuji, S. and Zheng, J.Y.: Visual Path Planning by a Mobile Robot, Proc. 10th IJCAI, pp.1127-1130, 1987.
- ⑤富田, 高橋: 画像のB-REPのためのアルゴリズム, 信学会パシフィック・学習情報, PRU86-87, 1987
- ⑥富田, 高橋: ステレオ画像の境界線表現に基づく対応, 信学論, J71-D, 6, 1988.
- ⑦富田, 杉本, 高橋: ステレオによるオクルージョンの検出, 第36回情報処会全大, 1988.
- ⑧Yachida, M.: 3D Data Acquisition by Multiple Views, Robotics Research, pp.11-18, MIT Press, Cambridge, Mass., 1986.
- ⑨Ayache, N. and Lustman, F.: Trinocular Stereovision: Recent Results, Proc. 10th IJCAI, pp.826-828, 1987.
- ⑩Waltz, D.: Generating Semantic Descriptions from Drawings of Scenes with Shadows, Winston, P. (Ed.), The Psychology of Computer Vision, pp.19-92, McGrawHill, New York, 1975.