

南カリフォルニア大学でのビジョン研究と動画解析について

(Computer Vision Research at U.S.C. and Motion Analysis)

安本 吉雄 (Yoshio Yasumoto)

松下電器産業株式会社 (Matsushita Electric Industrial Co., Ltd.)

Research activities on computer vision at U.S.C. Intelligent Systems Group headed by Prof. Nevatia are described based on the author's 2 years' experience with this group. Research environment, past establishments and the edge & corner detection are introduced. Robust motion parameter estimation with regularization is also described as an example of new projects of this group.

1. まえがき

著者は1983年8月から1985年8月まで、米国ロスアンジェルス南カリフォルニア大学(University of Southern California, 以下U.S.C.と略す)の電気工学科に客員研究員として滞在する機会を得た。USCの電気工学科は画像処理・デジタル信号処理等の分野で1970年代からよく知られているが、ビジョンの分野では比較的新しい。この分野ではネバティア準教授を中心とするグループがインテリジェント・システムズ・グループという名の元に従来の画像処理研究所(Image Processing Institute)から独立した。本稿では、当グループにおけるビジョン研究の現状について紹介するとともに、著者らが現地で行ったコーナー検出・動きのパラメータ抽出等の研究について報告する。

2. USC Intelligent Systems Group

当グループは、航空写真解析や三次元物体の表現を研究していたスタンフォード大学出身のネバティア準教授を中心とするグループで、他に2名の助教授とPh.Dの学生数名から構成されている。プライス助教授はC.M.U.でセグメンテーション等の研究をしていたが、U.S.C.においては航空写真解析、知識ベースの導入等に関心をもっている。メディアオニ助教授はPh.D取得後、著者が着任した1983年9月より助教授になった人で、Ph.Dの論文テーマはセグメントを用いたマッチング法、ステレオ視への応用であった。

研究用の計算機環境としては従来よりDECのKL-10(OSはTOPS20)を主として利用し、画像表示用にComtal社のVision 1を2台使用していた。著者の滞在中にDECのVAX11/750(UNIX)が2台導入され移行しつつあった。さらにSymbolics社の3640が3台導入されVAX11とリンクされて画像表示やLISPプログラムの高速実行用に使用されている。使用言語は従来のSAILよりC,LISPへと移行しつつある。TOPS20やUNIX上で使用するエディターはBMACSで、論文清書用にはScribeがよく使われていた。

当グループは従来画像処理研究所(Image Processing Institute)の一部門であったが、著者が着任した時点ではすでに別個に運営されていた。画像処理研究所は1970年代、米国の画像処理研究のメッカ的位置付けにあったが、初代所長プラット氏がVicom社を設立して出、2代目所長アンドリュース氏もComtal社副社長として辞職してからは目立った成

果は出ない。現在は3代目所長ソーチャク教授の元にデジタル信号処理のクング準教授、画像処理、ビジョンのチェラバ助教授、光学画像処理のジェンキンス助教授らとグループを構成し名称もSignal and Image Processing Instituteと変更した。

一方、計算機学科にはビーキ教授を所長とするロボティクス研究所(Robotics Institute)が存在しIntelligent Systems Groupとの関係も深い。ロボティクス研究所ではIBMの援助でマニピュレーターが数台導入され、当グループと共同使用している。なおネバティア準教授は計算機学科の併任教授でもある。

3. U. S. C. I. S. G. のビジョン研究

USCにおけるビジョン研究は米国の他大学と同様、国防総省DARPAの資金援助による画像理解プロジェクトの一員としてスタートした。

| | |
|-----------|---------------------|
| DARPA | basic I. U. Project |
| | Aerial Mapping |
| | Robotic Vision |
| DMA | Parallel Processing |
| | Mapping |
| | Building Analysis |
| DARPA/ALV | Motion Detection |

第1表 I. S. G. のプロジェクト

| |
|--|
| <ul style="list-style-type: none"> • M. I. T. • C. M. U. • Stanford Univ. • U. S. C • Univ. of Massachusetts • Rochester Univ. • Univ. of Maryland • Columbia Univ. • S R I |
|--|

第2表 I. U. プロジェクト・メンバー

1985年8月時点でのプロジェクトを項目別にまとめると第1表のようになる。またI. U. プロジェクトに参加の大学・研究機関は第2表のとおりである。

4. サブピクセルを利用したエッジ検出とコーナー検出

4. 1 Haralick(1) は画像内の小ウィンドウに対して連続関数を得るためのファセットモデルを提案した。連続関数として2次元直交チェビシェフ多項式を用い、ゼロクロッシングによるエッジ検出・コーナー検出を実現した。小ウィンドウW内の画像をImとすると

$$I_m(r, c) = \sum_{n=0}^{N-1} a_n P_n(r, c) \quad (1)$$

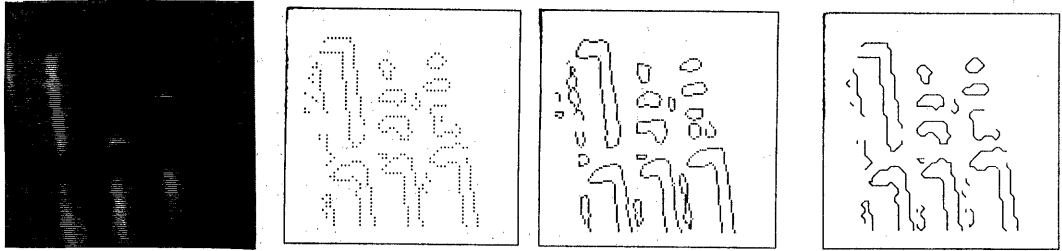
ここで $P_n(r, c)$ は2次元直交チェビシェフ多項式、Nはウィンドウの大きさを示す。

係数 a_n は

$$a_n = \frac{\sum_r \sum_c P_n(r, c) I_m(r, c)}{\sum_i \sum_j P_n^2(i, j)} \quad (2)$$

4. 2 U. S. C. のHuertasら(2)はHaralickのファセットモデルを用いてサブピクセル

ル精度のエッジ検出を実現した。第1図はその例で原画は64×64で8ビットである。従来のピクセル精度のゼロ・クロッシング法によるエッジ検出結果と、2倍の精度のサブピクセルを用いた結果を比較している。この例では128×128画素に拡大し9×9のラプラシアン・ガウシアンマウクとコンボリューションをとりエッジを求めた。結果をみると明らかなように、2倍の精度のサブピクセルを用いたエッジの方が通常のエッジよりも解像度の高い良好なものとなっている。特にエッジの連続性についてはすぐれた結果となっているが、これはコーナー検出のための曲線追跡に重要である。



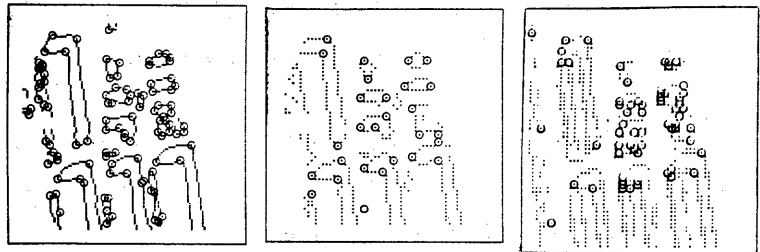
第1図

第2図

4. 3 第2図はU. S. C. で用いられている標準方式であるP. S. ファイルを示している。これは Nevatia-Babu線型特徴抽出法(3) で用いられた手法で、エッジの連結性を画像ファイルで登録したものである。ゼロ・クロッシングで求めたエッジの場合細線化の必要はないが、エッジの方向は予測法等で求めなければならない。

4. 4 Zunigaら(4) はファセットモデルを用いたコーナー検出を提案している。これは連続関数を元にして各点における曲率と方向を計算したものをを用いている。これ以外にも従来より種々のコーナー検出法が提案されているが、これらは同一だとする報告もある。(5) 一方簡易な方法

としてはk-curvature 法が知られているが、いずれも曲率を計算するもので、画像の特徴を忠実に表現するには不足している。Langridge (6) は3



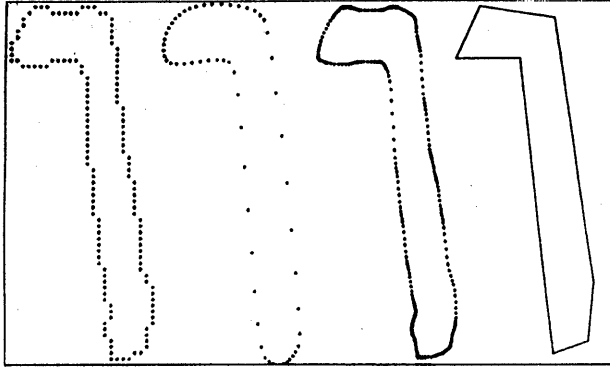
第3図

次多項式を曲線にあてはめてくり返し法により不連続点を見付ける手法を報告している。

著者ら(7) はサブピクセル精度で求めたエッジ曲線にB-スプラインをあてはめ、各点を元の位置からずらせてその変位を測定した。各点における曲率もあわせて計算し、変位と曲率の大きさよりコーナーを判定した。この方法によると各点の位置情報だけでくり返し法を用いることなく、コーナー性を判定できると同時にエッジ曲線をB-スプラインで表現することができる。この表現法はコーナーを不連続点とするB-スプラインを用いているので、ビジョン関係への応用が期待できる。

第3図に各種コーナー検出結果を比較する。左より本方法、3-curvature 法、Zunigaら

による方法である。第4図は各種曲線表現法の比較で、左より原図、コーナーを含まないスプライン表現、本方法、Nevatia-Babu線型特徴抽出法による直線表現である。



第4図

5. 正規化法を用いた 動きのパラメータ抽出

5.1 近年剛体の動きのパラメータを求める研究が盛んに行なわれている。初期の研究は動きに一定の制限を設け、非線形方程式に帰着させる方法が多かった。最近Tsai, Huang (8) と Longuet-Higgins (9) は独立に、線型方程式を導くことに成功した。実画像に

応用し得る観点からみると、入力カメラや量子化誤差、特徴点抽出に伴う誤差等、すべての誤差を吸収してかつ安定な解を得られるものでなければならない。この意味で、Tsaiらの方法は計算は容易だが実用的でない。それは彼らも認めているところで、入力データの誤差が1%あると、求めたパラメータの値が60%変化するという。一方Bruss と Horn (10) は実測した画面上の動きと計算した動きの差を最小にするために、最小二乗法を用い、ノイズに強い手法を導いた。同様にAdiv (11) は平行移動の各パラメータを求めるために最小二乗誤差関数を導入し、各パラメータに対して線型方程式を解き回転パラメータを求めた。著者ら (12) は動きのパラメータを求める問題を「明白でない逆問題」と考え、正規化項を誤差関数に導入した。

5.2 実際的な問題の多くは明白な解のない問題である。「明白な解が存在する問題」とは定義により、解がただ一つ存在し解と入力データの対応が連続である問題である。定義より、動きのパラメータ抽出は「明白な解の存在しない問題」であり、また逆問題でもある。

正規化法 (Regularization) はこれらの問題に対処するため、解の存在領域に一定の制限を設けて解決しようとする方法で、20年前よりソ連の数学者ら (13) によって発展させられ最近Poggio (14) が画像への応用を提案した。

一般に逆問題に対する解法手順は次のようになる。まず入力を Y とすると、今 $Az=Y$ なる解 z を求めようとする。そのためノルム $\| \cdot \|$ を定義し、さらに安定化関数 Pz を導入する。ここで問題は、 $\|Az-Y\|^2 + \lambda \|Pz\|^2$ を最小にする関数を求めることに帰着する。

第1項は入力データと解の距離を表現し、第2項は正規化の度合を示している。

動きのパラメータを求める問題においては、平行移動と回転の各パラメータが入力の変化に対して滑らかであってほしい。しかし平行移動の各パラメータは互いの比のみしか得られないので、これらを正規化項に導入することは難しい。ここでは回転パラメータ Ω のベクトル長を安定化関数 Pz とした。すなわち $Pz = \sqrt{\Omega_x^2 + \Omega_y^2 + \Omega_z^2}$ 。

5.3 3次元空間にある剛体の動きに関する基本式は、

$$(x' \ y' \ z') = R (x \ y \ z) + T \quad (3)$$

ここで (x, y, z) は剛体上点P の動き前の空間座標で、 (x', y', z') は点P の動き後の空間座標、またR は 3×3 の正規行列でその各要素は回転角 θ と回転軸の方向ベクトル (n_1, n_2, n_3) で表わされる。T は平行移動のベクトルで $T = (T_x, T_y, T_z)^T$ である。

ここで回転角 θ が小さいと仮定すると、行列R は

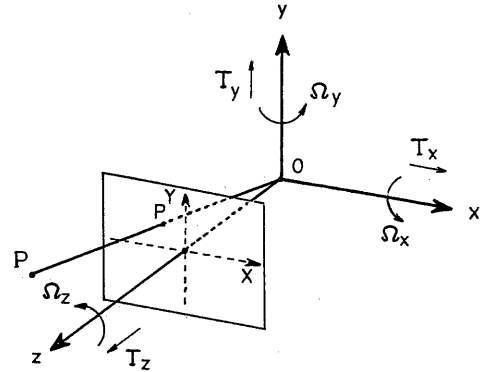
$$R = \begin{bmatrix} 1 & -\Omega_z & \Omega_y \\ \Omega_z & 1 & -\Omega_x \\ -\Omega_y & \Omega_x & 1 \end{bmatrix} \quad (4)$$

で表わされる。 $\Omega_x, \Omega_y, \Omega_z$ はそれぞれX, Y, Z 軸のまわりの回転角を示し、第5図のようになる。

空間座標 (x, y, z) と画像面座標 (X, Y) の関係はよく知られているように、

$$X = f x / Z, \quad X' = f x' / Z' \quad (5-1)$$

$$Y = f y / Z, \quad Y' = f y' / Z' \quad (5-2)$$



第5図

ここで f はカメラの焦点距離を、 $(X, Y), (X', Y')$ は動き前後の画面上の点P, P'を示す。

いま連続する2画面間の画像面上の点の移動ベクトル $(X' - X, Y' - Y)$ を (α, β) と置き換え、式(3)-(5)を用いてR, T 及びZ の関数として次のように表現する。

$$\alpha = \frac{-\Omega_x X Y + \Omega_y (f^2 + X^2) - \Omega_z f Y + (T_x f - T_z X) f / z}{f + \Omega_x Y - \Omega_y X + T_z f / z} \quad (6-1)$$

$$\beta = \frac{-\Omega_x (f^2 + Y^2) + \Omega_z f X + (T_x f - T_z Y) f / z}{f + \Omega_x Y - \Omega_y X + T_z f / z} \quad (6-2)$$

ここで動きを独立な2つの要素、回転と平行移動に分割し整理して書き直すと、

$$\alpha = \alpha_R + \alpha_T, \quad \beta = \beta_R + \beta_T \quad (7)$$

ただし、

$$\alpha_R = -\Omega_x \frac{X' Y}{f} + \Omega_y \left(f + \frac{X X'}{f} \right) - \Omega_z Y \quad (8-1)$$

$$\beta_R = -\Omega_x \left(f + \frac{Y Y'}{f} \right) + \Omega_y \frac{X Y'}{f} + \Omega_z X \quad (8-2)$$

$$\alpha_T = \frac{T_x f - T_z X'}{Z} \quad (8-3)$$

$$\beta_T = \frac{T_y f - T_z Y'}{Z} \quad (8-4)$$

問題は観測した画面上の各点の移動ベクトル (α, β) から動きのパラメータ $\Omega_x, \Omega_y, \Omega_z$, 及びT を求めることである。前節で説明したように、最小にすべき誤差関数として観測した移動ベクトルと計算した値との最小二乗誤差及び正規化項を含むものを選んだ。

$$\sum_i [(\alpha_i - \alpha_{Ri} - \alpha_{Ti})^2 + (\beta_i - \beta_{Ri} - \beta_{Ti})^2] + \lambda (\Omega_x^2 + \Omega_y^2 + \Omega_z^2) \quad (9)$$

ここで α_i, β_i は i 番目の点の移動ベクトルで、 n 個の点について評価する。

平行移動のベクトル T と各点の深さ方向の値 Z_i は絶対値が得られないので、ベクトル T の長さを $r (r \neq 0)$ と仮定し、平行移動単位ベクトル U と各点の相対的奥行き値 d_i を導入すると、

$$(U_x, U_y, U_z) = (T_x, T_y, T_z) / r, \quad d_i = r / Z_i \quad (i=1 \dots n)$$

これらを用いて式(9)を書き直すと、

$$\sum_i [(\alpha_i - \alpha_{R_i} - \alpha_{U_i} d_i)^2 + (\beta_i - \beta_{R_i} - \beta_{U_i} d_i)^2] + \lambda (\Omega_x^2 + \Omega_y^2 + \Omega_z^2) \quad (10)$$

$$\text{ここで } \alpha_{U_i} = U_x f - U_z X_i = \alpha T_i / d_i \quad \beta_{U_i} = U_y f - U_z Y_i = \beta T_i / d_i \quad (11)$$

このようにして問題は回転角 $\Omega_x, \Omega_y, \Omega_z$ と平行移動単位ベクトル U 及び各点の相対的奥行き d_i を求めることになった。次に式(10)の d_i に関する導関数を用いて d_i を消去する。導関数を 0 と置くことによって、分母が 0 でない場合は、

$$d_i = \frac{(\alpha_i - \alpha_{R_i}) \alpha_{U_i} + (\beta_i - \beta_{R_i}) \beta_{U_i}}{\alpha_{U_i}^2 + \beta_{U_i}^2} \quad (12)$$

となる。ここで剛体上の各点はカメラの前方に存在するから、 d_i は常に正数でなければならぬ。誤差関数(10)を評価する時に式(12)の分子 d_i が常に正であるかどうかを確認する必要がある。このようにして式(10)の誤差関数に含まれる各点の誤差関数 E_i は式(12)を用いて

$$E_i = \frac{[(\alpha_i - \alpha_{R_i}) \beta_{U_i} - (\beta_i - \beta_{R_i}) \alpha_{U_i}]^2}{\alpha_{U_i}^2 + \beta_{U_i}^2} \quad \delta_i > 0 \quad (13)$$

と書ける。また最小にすべき誤差関数 $E(U, \Omega)$ は動きのパラメーターだけの関数として

$$E(U, \Omega) = \sum E_i + \lambda (\Omega_x^2 + \Omega_y^2 + \Omega_z^2) \quad (14)$$

となる。ただしこの関数はすべての δ_i が正である時のみ定義される。

誤差関数を評価する方法として、すべての可能性あるパラメーターの値についてそれぞれ誤差関数の値を求めることにする。幸い平行移動単位ベクトルを採用しているので、半径が 1 の球面上を探索だけでよく、式(14)は球面上の点対称の点に対して同じ値をとるので、半球面上を検索すると同時に条件 $\delta_i > 0$ を両方の点に対して確認するだけでよい。半球面上の一点が与えられると、その点において各回転パラメーター $\Omega_x, \Omega_y, \Omega_z$ に関する偏微分をとり 0 と置くことによって、各回転パラメーターに対する 3 元一次の連立方程式が得られる。この方程式より各回転パラメーターが求められ、半球面上の一点の誤差関数の値が得られる。

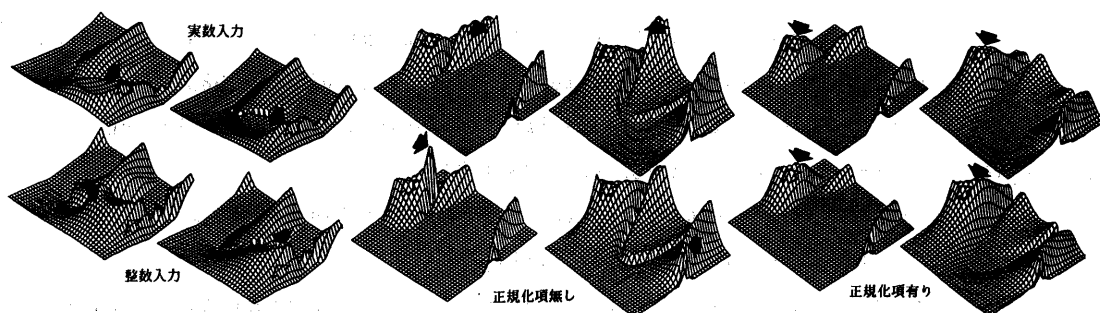
5. 4 半球面上を検索し平行移動単位ベクトルに対し誤差関数 E の値を求める上で、次の置換を行なった。

$$U_x = \sin \phi \cos \theta, \quad U_y = \sin \phi \sin \theta, \quad U_z = \cos \phi \quad (15)$$

ここで $0 \leq \phi \leq 90^\circ, 0 \leq \theta \leq 360^\circ$ で半球面上の各点を表示できることになる。次節に示す誤差関数は θ と ϕ を 2 軸として表示してあり、誤差関数の値は正負を逆にしてあるので最大値が誤差関数の最小点を意味する。また白い矢印が正解値を、黒い矢印が計算値を示す。今回使用したデータはすべて計算機で生成した画像であり、あらかじめ与えたパラ

メーターに基づいて画面上の移動ベクトルを計算して使用した。

5. 5 ここでは2つの計算例を示す。第6図は画像の入力精度によって得られたパラメーターがどう変化するか、また正規化法によってそれがどの程度改善されるかを示すものである。ここで用いたデータは $\Omega_x=1.0^\circ$ 、 $\Omega_y=0^\circ$ 、 $\Omega_z=-1.0^\circ$ と回転角が仮定のとおり小さく、平行移動単位ベクトルは $U_x=U_y=U_z/2$ である。画面上での移動ベクトルは大きさが10~30ピクセル程度で通常カメラで移動する剛体を画面の周辺にとらえた場合に相当する。結果をみると実数入力の場合は正規化項を導入しなくても、計算値が正確に近い値を示すが、整数入力の場合は正規化項がない場合は正解から離れた値を示す。ここで実数入力とは移動ベクトルのXY座標として実数を選択し、整数入力は整数を選択したもので、通常の実画像の場合は整数しか得られない。整数入力の場合は正規化項の効果が大いといえるが、他の数多くの例でも確認された。



第6図

第7図

第7図はカメラが一定の動きのパラメーターで動いたときに静止している剛体の画面上での動きを仮定したものである。この仮定はマニピュレーターに取付けられたカメラが一定のパラメーターで動いた場合に得られる連続画像に対応している。用いたデータは $\Omega_x=1.0^\circ$ 、 $\Omega_y=0^\circ$ 、 $\Omega_z=-2.0^\circ$ 、 $U_x=-2U_y=U_z$ なる動きのパラメーターで作成し、整数入力とした。整数にしたための誤差は各座標の絶対値によっても違うが、0.2~1.2%の範囲にある。また正規化項の係数 λ はこの場合 $\lambda=5.0$ に設定した。結果をみると正規化項の含まない誤差関数は各フレーム間で得られたパラメーターが大きくばらついているが、正規化項を含む場合は常に正解値に近い値が得られている。

5. 6 正規化項の効果としてノイズの存在する場合、解の安定性に大きく寄与することがわかった。これは実際の動画解析において適用し得る有効な方法と考えられる。ただし画像とノイズの有効なモデルを導出し、最適な λ の値を選択する必要は残されている。さらにここで示された誤差関数の形が、かなりの領域にわたって平坦すぎるので、極値を求めるのを困難にしている。これは今まで用いられた各種の制約条件が充分でないことを示しているものであり、別の観点から解決をはかる必要があると思われる。この解決法の一つは画面上での動きの加速度を測ることであり、線対応を考えることが必要となろう。

6. 謝辞

南カリフォルニア大学での研究でお世話になったネバティア教授及びメディオニ助教授に感謝します。著者が南カリフォルニア大学に滞在する上でお世話になった京都大学の長尾教授及び東北大学の松山助教授に謝意を表します。また海外留学の機会を与えて下さった松下電器産業株式会社・本社及びテレビ事業部の関係者各位特にテレビ事業部開発部山本部長に感謝します。

<参考文献>

- 1) R. M. Haralick, "Digital Step Edges from Zero Crossing of Second Derivatives," IEEE Trans. on PAMI-6 p58-68(1984)
- 2) A. Huertas & G. Medioni, "Edge Detection with Subpixel Precision," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, p633-636(1984)
- 3) R. Nevatia and K. R. Babu, "Linear Feature Extraction and Description," Computer Graphics and Image Processing 13, p257-269(1980)
- 4) O. A. Zuniga and R. M. Haralick, "Corner Detection Using the Facet Model," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, p30-37(1983)
- 5) M. A. Shah and R. Jain, "Detecting Time-Varying Corners," Computer Vision, Graphics, and Image Processing 28, p345-355(1984)
- 6) D. J. Langridge, "Curve Encoding and the Detection of Discontinuities," Computer Graphics and Image Processing 20, p58-71(1982)
- 7) G. Medioni and Y. Yasumoto, "Corner Detection and Curve Representation Using Cubic B-Splines," Proc. IEEE Conf. on Robotics and Automation (to be published)
- 8) R. Y. Tsai and T. S. Huang, "Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," IEEE Trans. on PAMI 6, p13-26(1984)
- 9) H. C. Longuet-Higgins, "A Computer Algorithm for Reconstructing a Scene from Two Projections," Nature 293, p133-135(1981)
- 10) A. R. Bruss and B. K. P. Horn, "Passive Navigation" MIT A. I. Memo 662(1981)
- 11) G. Adiv, "Determining 3-D Motion and Structure from Optical Flow Generated by Several Objects," Proc. of Image Understanding Workshop(1984)
- 12) Y. Yasumoto and G. Medioni, "Experiments in Estimation of 3-D Motion Parameters from A Sequence of Image Frames" Proc. of Computer Vision and Pattern Recognition p89-94(1985)
- 13) A. N. Tikhonov, "The Regularization of Ill-Posed Problems" Dokl. Akad. Nau. SSR 153(1), p149-152(1963)
- 14) T. Poggio and V. Torre, "Ill-Posed Problems and Regularization Analysis in Early Vision," Proc. of Image Understanding Workshop(1984)