

解剖学的データベースを用いた 単眼陰影画像からの3次元顔モデル復元

吉城 香織[†] 齋藤 英雄[†] 持丸 正明^{††}

[†] 慶應義塾大学工学部 情報工学科

^{††} 産総研デジタルヒューマン研究センター

E-mail: †{yoshiki,saito}@ozawa.ics.keio.ac.jp

あらまし 本稿では、1枚の正面顔画像から手軽に顔形状復元を行う手法を提案する。解剖学的特徴を多く含むデータベースを主成分分析し、寄与率の高い主成分のみを形状の推定に用いることで、少数の形状パラメータから復元を行う。形状を復元すると同時に、顔形状の特徴を解剖学的に表すデータを直接求めることができる利点を持つ。比較は画像同士で行い、データベース中のモデルから作成した画像と入力画像とを比較する。比較の際に陰影情報や特徴点の適切な評価値を与えることで、評価値が小さくなる方向へモデルの位置・姿勢、形状パラメータを更新し、最適化により顔形状を復元する。実験の結果、平均誤差2.62mmの復元精度を実現した。またレンジデータに近い精度での復元が行われたことを示し、眼鏡のオーダーメイド等のアプリケーションへ利用できる可能性を示した。

Reconstruction of 3D Face Model from Single Shading Image Based on Anatomical Database

Kaori YOSHIKI[†], Hideo SAITO[†], and Masaaki MOCHIMARU^{††}

[†] Department of Information and Computer Science, Keio University

^{††} Digital Human RC, AIST

E-mail: †{yoshiki,saito}@ozawa.ics.keio.ac.jp

Abstract We propose a method to reconstruct 3D face model from a single face-frontal shading image based on a database of the original 3D face shapes which are created according to anatomical basis of the human head. We reconstruct the 3D shape of the face by estimating small number of eigenvalues taken by Principal Component Analysis(PCA) of the anatomical database of the head, rather than directly recovering the shape from the input shading image. The eigenvalues are estimated based on the optimization of the error value computed by comparing shading information of input image to that of model image created from the shape represented by the eigenvalue. The reconstructed shape for a standard head model with the average shape of the databases provides an error evaluation, which is 2.62mm in average. This is sufficiently accurate for the future applications, such as on-line order-made production of glasses, and so on.

1. 序 章

従来、コンピュータービジョンの分野において、物体を3次元形状で表現する研究が広く行われてきた。レンジデータ等の形状測定装置が開発され、大規模空間から人間の顔形状のような微細な形状に至るまで、精密な形状情報を手にすることが可能になった。このような精密な測定装置を用いて、物体の精密な形状を直接測定しようというアプローチが盛んに行なわれている。

こうして精密に計測された顔形状情報を、実際の販売の場や実験の場で利用しようとする要望が出てきている。例として、眼鏡のオーダーメイド、ガスマスクの製造等が挙げられる。しかし、特殊な撮影環境が必要であったり、測定装置が大規模となる、精密な測定を行うには測定時間がかかる、等の問題から3次元形状を実際の販売の場で利用することは困難であるとされてきた。こうした問題を解決するために、その場で直接形状を計測するのではなく、手軽に撮影可能なデジタルカメラやビデオカメラ等の画像から、ユーザの顔形状を復元しようとする

$$T^T = \{X_1, Y_1, Z_1, X_2, Y_2, Z_2, \dots, X_{480}, Y_{480}, Z_{480}\}$$

●メッシュデータ 830 個

2.2 主成分分析

本手法では、顔形状を低次元の形状パラメータで表すため主成分分析を用いる。データベースの \$m\$ 人の頂点データ \$(T_1, \dots, T_m)\$ で表される行列 \$M\$ (式 1) を主成分分析し、主成分 \$P\$ を得る。主成分得点を形状パラメータとし、これを变化させることで形状を更新し最適な復元形状を推定する。

$$M = [T_1, T_2, \dots, T_m]^T \quad (1)$$

\$n\$ 個の頂点をもった顔形状のうちの一つを \$T = (t_1, t_2, \dots, t_n)\$ とする。また、平均顔形状を \$\bar{T} = (\bar{t}_1, \bar{t}_2, \dots, \bar{t}_n)\$ とする。ここで、第 \$q\$ 主成分 \$p_q = (r_1, r_2, \dots, r_n)\$ に対応する主成分得点 \$s_q\$ は以下のように表される。

$$s_q = \sum_{i=1}^n r_i (t_i - \bar{t}_i) \quad (2)$$

第 \$n\$ 主成分までの主成分得点を \$S_n = (s_1, s_2, \dots, s_n)\$ とし、第 1 から第 \$n\$ 主成分を \$P_n = (p_1, p_2, \dots, p_n)\$ とおくと、

$$S_n = P_n^T (T - \bar{T}) \quad (3)$$

と書ける。式 3 を変形させると、\$P^{-1} = P^T\$ より、最終的な復元形状復元形状 \$T\$ は式 4 で得られる。

$$T = \bar{T} + P_n \times S_n \quad (4)$$

2.3 復元の流れ

本手法では、入力画像と、データベースのモデルから作成した画像との比較を行う。複数の画像を用いないため、画像中の陰影情報が復元の際に重要な情報として扱われる。よって、いかに入力画像と比較可能のようにモデルの陰影情報を決定するかが重要となる。

また、入力画像が完全に正面を向いているとは限らないため、モデルから画像を作成する際、形状モデルの画像平面に対する位置・姿勢の 6 自由度を推定することが求められる。求めた位置・姿勢で主成分得点 \$S_n\$ で代表される \$n\$ 自由度を持つ形状パラメータを推定し、顔形状を復元する。よって復元形状を得るには、入力画像と一致するモデルの位置・姿勢、形状の \$(n+6)\$ 自由度を持つパラメータを推定する必要がある。このパラメータの推定の際、評価値を与え、評価値が最初となるよう最適化を行う。そのため適切な評価値を与えることが重要となる。

本章では 2.3.1 節でモデルから画像を作成する手法を述べ、2.3.2 節で本手法で用いた評価値について述べる。2.3.3 節で最適化処理の流れを述べる。

2.3.1 モデル画像作成

本手法ではモデルから画像を作成する際、データベースの頂点・パッチデータを用いて、透視投影によりモデル画像を作成する (図 2)。また顔表面のシェーディングモデルとして、Lambert の余弦則による拡散反射モデルと Phong の鏡面反射モデルを組み合わせたものを用いた。ほぼ全ての物体は拡散反

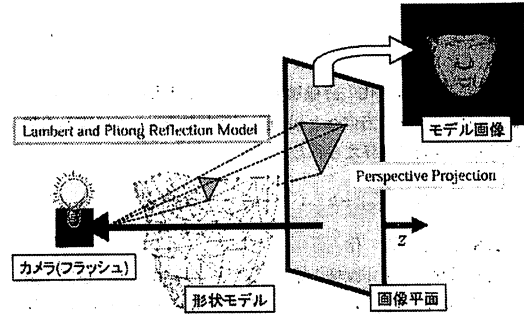


図 2 モデル画像作成法

射モデル鏡面反射モデルを組み合わせたもので表現される。これらのモデル中の反射係数などのパラメータを変えることによって、様々な材質感を表現することができる。被写体に対し光源は正面に存在し、平行照射であると仮定する。また本研究では顔の表面状態や光源の強さは既知であるものとする。

\$I_{in}\$ を入射光の強度、\$k_a, k_b\$ はそれぞれ顔表面の拡散反射係数、鏡面反射係数、\$\alpha\$ を照明の入射方向と顔表面の法線方向のなす角度、\$\beta\$ を照明の反射方向と視線のなす角度、\$u\$ を鏡面反射係数とすると、モデル画像における各画素の明度値 \$I\$ は式 5 のように表される。

$$I = I_{in} \times (k_a \cos \alpha + k_b \cos^u \beta) \quad (5)$$

2.3.2 評価関数

後述の 2 つの評価関数を用いて評価値を計算し、最適化に用いた。

(1) 正規化相関 \$C_e\$

画像全体での正規化相関を求め、評価関数として与える。正規化相関は画像間の陰影情報の類似度を表す値であり、テンプレートマッチングにおける類似度を表す評価尺度として頻繁に用いられる。対応する画素値の平均と標準偏差によって正規化されており、画素値の線形変換に対して一定の評価値を返すので、照明変動など環境変化の影響を受けにくい。\$N\$ を全画素数、\$I_i, \bar{I}_i\$ を入力画像の明度値とその画像全体での平均、\$I_m, \bar{I}_m\$ をモデル画像の明度値とその画像全体での平均とすると、正規化相関 \$C_e\$ は式 6 で表され、\$-1 \sim 1\$ の値を取る。値が 1 に近いほど類似度は高くなる。

$$C_e = \frac{\sum_{i=0}^{N-1} (I_i(i) - \bar{I}_i) \times (I_m(i) - \bar{I}_m)}{\sqrt{\sum_{i=0}^{N-1} (I_i(i) - \bar{I}_i)^2} \sqrt{\sum_{i=0}^{N-1} (I_m(i) - \bar{I}_m)^2}} \quad (6)$$

(2) 特徴点誤差 \$C_f\$

ユーザが入力画像上で左右の目尻、左右の口の端点の 4 点を指定する。この 4 点の画像上でのユークリッド距離の二乗誤差を評価関数として与える。この距離が小さくなる方向へ最適化が行なわれる。入力画像上での特徴点位置を \$f_i(i) (i = 1, \dots, 4)\$、モデル画像上での特徴点位置を \$f_m(i) (i = 1, \dots, 4)\$ とすると、評価関数 \$C_f\$ は式 7 で表される。

$$C_f = \sum_{i=1}^4 |f_i(i) - f_m(i)| \quad (7)$$

最終的に、最適化の評価値 C は式 8 で与えられる。この評価値が小さくなる方向へと最適化が行われる。 a, b は重み係数であり、経験的に与えられる。

$$C = a \times (1 - C_e) + b \times C_f \quad (8)$$

2.3.3 最適化

入力画像とモデル画像の陰影情報が一致するよう、位置・姿勢、形状の $(n+6)$ 自由度のパラメータの最適化を行う。滑降シンプレックス法を用いて評価値が最小となる際の形状パラメータを求め、復元形状とした。最適化の流れを以下にまとめる。

(1) データベースの初期形状からモデル画像を作成する。特徴点 4 点を指定し、入力画像に対する位置・姿勢を大まかに与え作成する。

(2) モデルの 3 次元位置・姿勢パラメータの 6 自由度を推定するため、入力画像とモデル画像を比較し、評価値を算出する。

(3) 評価値が小さくなる方向へ位置・姿勢を変化させモデル画像を更新する。

(4) (2), (3) を繰り返し、最適化によって評価値が収束するモデルの適切な位置・姿勢を決定する。

(5) 主成分得点の上位 n 次元で表される形状パラメータを推定するため、入力画像と (4) で求めた位置・姿勢におけるモデル画像を比較し、評価値を算出する。

(6) 評価値が小さくなる方向へ形状を変化させモデル画像を更新する。

(7) (4), (5) を繰り返し、最適化によって評価値が収束するモデルの形状パラメータを決定する。

(8) (6) から、最終的な入力画像の復元顔形状を得る。

3. 実験

本手法の復元精度を評価するため、2 種類の実験を行った。実物大の形状の既知な模型を撮影し入力画像として復元を行い、正解形状との頂点誤差を求めることで評価を行った。また実際の人物を撮影し、入力画像として復元を行った。対象人物のレンジデータを撮影し、正解形状として画像上で比較を行い、その精度を検証した。

両実験とも上位 20 主成分を復元に用い、形状パラメータとして 20 自由度を持つ主成分得点を変化させ、最適化を行うことで復元形状を求めた。また特徴点として口の両端と左右の目の 4 点を入力画像上で与え、評価値として用いた。

3.1 形状の既知な模型を用いた実験

形状の復元精度を評価するため、データベースの平均顔の形状を持つ実物大の模型を撮影し、入力画像として本手法により復元を行った。サイズは 256×256 とし、頭部領域を切り抜いてグレースケール化され陰影情報として与える。通常は初期形状にデータベースの平均顔形状を用いるが、模型が平均形状を持つため、初期モデル形状は平均形状ではなくデータベース中

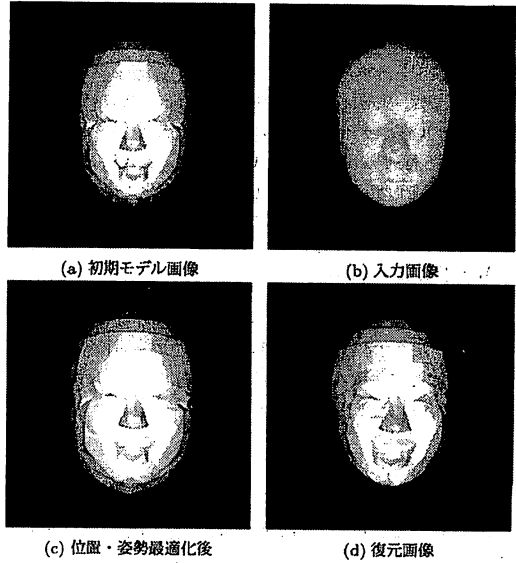


図 3 結果画像

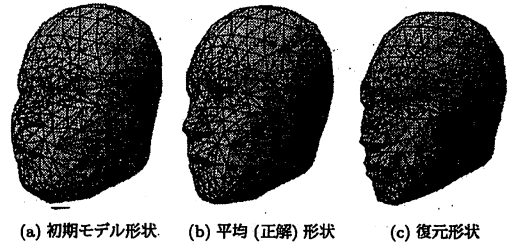


図 4 別視点から見た復元形状

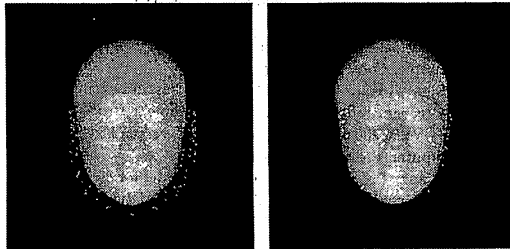
の任意の形状とした。

結果画像とその別視点からの画像を図 3 に載せる。入力画像に対し、初期モデル画像は位置・姿勢、形状共にずれている。(b) の入力画像はほぼ正面から撮影されたが、(a) の初期モデル画像と比較して上方方向を向いている。これを補正するために位置姿勢の最適化を行った。(c) の位置・姿勢最適化後の画像を見ると、(a) が初期モデル画像より上方方向を向いており、入力画像とほぼ等しい位置・姿勢をしているとわかる。また復元後の輪郭が、復元前に比べ正解形状にフィットしていることがわかる。

データベース中の任意のモデル 6 体を初期モデルとし、それぞれに対して本手法を適用し、復元形状を求めた。復元前形状、復元後形状に対し正解形状との頂点誤差を算出して表 1 示す。誤差は顔部分の 260 頂点に対して算出し、平均・分散を求めた。復元前の形状では誤差が平均 4.30mm なのに対し、復元後では 2.63mm と減少が見られた。Sample2 のように復元前から誤差が 2.19mm と比較的小さい形状に対しても誤差が減少していることがわかる。また分散が全体的に小さくなっており、偏り無く形状が復元されていることがわかる。

表 1 頂点誤差の比較 (mm)

Sample	平均		分散	
	復元前	復元後	復元前	復元後
1	7.65	3.65	39.37	9.33
2	2.19	2.11	2.21	2.03
3	3.85	2.40	6.52	3.80
4	3.01	2.12	5.35	2.24
5	4.67	2.90	11.24	5.68
6	4.43	2.57	12.16	3.65
平均	4.30	2.63	-	-



(a) 復元前 (b) 復元後

図 5 正解形状との頂点誤差

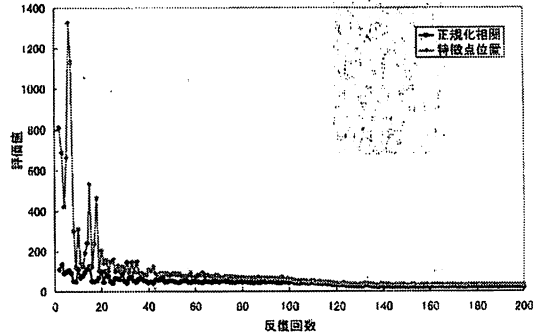
図 5 に, Sample1 の顔部分における誤差の分布を示す。位置・姿勢最適化で求めた位置・姿勢において, 復元前, 復元後の頂点を入力画像へ投影する。頂点を正解形状との誤差の大きさで色分けし, 赤が正解形状との誤差 4mm 以上, 黄色が 3mm 以上~4mm 未満, 青が 3mm 以下の頂点を示す。復元前の頂点誤差を図 5(a) に, 復元後の頂点誤差を図 5(b) に示す。(a) では頂点誤差が顔全体に広がっているが, (b) では全体的に誤差が減少して、正解形状へと近づいていることがわかる。

復元後を見ると, 輪郭部分に多く 4mm 以上の誤差が分布しているが, これは頂点が輪郭に沿って大きく動いてもモデル画像上での変化が見られないため, 誤差が生じたと考えられる。

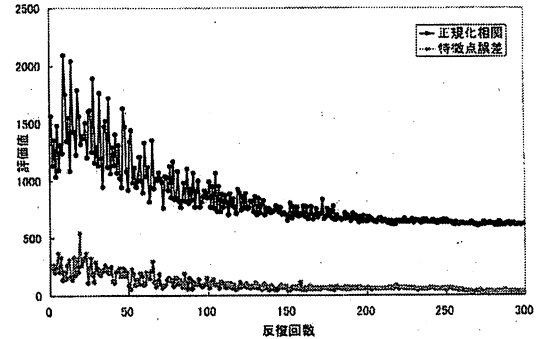
また, 最適化時の評価値の推移を図 6 に載せる。図 6(a) は位置・姿勢最適化時, 図 6(b) は形状最適化時の推移を示す。横軸は最適化反復回数であり, 縦軸は与えられた評価値である。今回, 陰影の類似度を示す正規化相関と, 左右の目尻, 口の両端の 4 点の特徴点誤差を評価値として与えた。この評価値が下れば, 設定した評価関数において, 最適化が効率よく行われていることが示される。正規化相関は逆数として与え, この値が減少するほど類似度が増加することを表す。この逆数の値, 特徴点誤差共に最適化が進むにつれ減少し収束している。従って, 設定した評価関数が位置・姿勢, また形状パラメータを推定する目的において, 適切であったことを示した。

3.2 実画像実験

実際の人物を撮影し, 入力画像として形状を復元した。対象人物の正解形状が未知であるため, 頭部形状をレンジスキャナで測定したレンジデータとの比較を行い, 正確に復元が行われているか評価した。模型実験同様入力画像のサイズは 256 × 256 であり, モデルに無い耳, 首領域を切り抜き, グレースケール



(a) 位置・姿勢最適化時



(b) 形状最適化時

図 6 評価値の推移



(a) 初期モデル画像 (b) 入力画像 (c) 復元画像

図 7 結果画像

として与える。初期モデル形状はデータベース中の平均顔形状とした。入力顔画像には頭髮が含まれているため, 頭頂部, 後頭部の形状は画像からは推定不可能である。従って本手法では, 復元モデルから頭頂部, 後頭部を除去した部位, つまり顔面部のみで画像を作成し比較を行った。

本手法を適用した結果画像を図 7 載せる。(a) の初期モデル画像では入力画像に比べ輪郭が小さいが, (c) の復元画像では入力画像に沿った輪郭が得られている。画像上では入力画像への最適化が適切に行われていることがわかる。

次に, 復元された 3 次元形状の比較を行う。対象人物の顔をレンジファインダで撮影し, レンジデータを得た。本実験で用いたレンジファインダは, スペースビジョン社の Cartesia FACE SYSTEM [14] である。このレンジスキャナを使用することで, 対象人物の顔形状の 3 次元点群が得られる。このレンジデータ

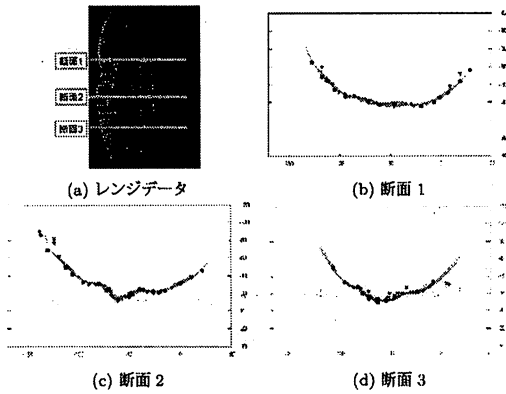


図8 レンジデータとの比較

を形状の正解形状として図 8(a) に示した。また図 8(a) における断面を上方からプロットしたものを図 8(b)(c)(d) に示す。復元後、復元前形状と正解形状との比較を行った。

赤が正解形状、紫が復元前、緑が復元後形状を 3 次元空間にプロットしたものである。復元前形状が正解形状に対して内側に入っており、正解形状より小さいのに対し、復元後の形状は正解形状に沿った形状となっている。全ての断面に対し、復元前の形状に比べ、復元後の形状が正解形状に近付いていることが見て取れる。以上のことから、本手法が精密に測定されたレンジデータに近い精度での形状の復元が可能であることを示した。

4. 結 論

本研究では、解剖学的顔形状データベースを用いて、正面方向から撮影された単眼陰影顔画像から顔形状モデルを復元する手法を提案した。

本手法では形状データベースを主成分分析して対象の顔形状を少ない次元のパラメータで表現し、このパラメータを入力した陰影画像から推定する。このため、陰影画像から対象の顔形状を直接復元する場合に比べて、より安定に形状復元を行うことが可能になる。

実験においては、復元前と復元後の形状を比較し誤差が減少していることを示した。また本手法がレンジデータに近い精度での形状復元を行っていることを示した。

また、解剖学的データベースを用いることで、形状を復元すると同時に頭部形状の特徴を解剖学的に表すデータを直接求めることができるという利点を持つ。このため、オンラインでの眼鏡のオーダーメイド、自動車事故のシミュレーション等のアプリケーションへの応用が期待できる。

文 献

- [1] Amit K. Roy Chowdhury, Rama Chellappa: "Face reconstruction from monocular video using uncertainty analysis and a generic model". Computer Vision and Image Understanding, Volume 91, Issues 1-2, July-August 2003, pp. 188-213
- [2] "Towards robust face recognition from multiple views".

- IEEE International Conference on Multimedia and Expo (ICME) 2004
- [3] Farhan Riaz, Asim Munawar, Arslan Mehboob, Wajahat Baig: "A New Algorithm for Reconstruction of 3-D Coordinates from a 2-D Image Using a Projected Pattern". IEEE INMIC, 2004
- [4] Ruo Zhang, Ping-Sing Tai, James Edwin Cryer, Mubarak Sha: "Shape from shading: a survey". IEEE Trans. Pattern Anal. Mach. Intell. 21 (8) (1999) 690-706.
- [5] J. Atick, P. Griffin, N. Redlich: "Statistical approach to shape from shading: reconstruction of three dimensional face surfaces from single two dimensional image". Neural Comput. 8 (1996) 1321.1340.
- [6] Wenyi Zhao, Rama Chellappa: "SFS based view synthesis for robust face recognition". Proceedings of the Fourth International Conference on Face and Gesture Recognition, Grenoble, France, 2000, pp. 285-292.
- [7] T. Sim, T. Kanade: "Combining models and exemplars for face recognition: an illuminating example". Proceedings of the CVPR 2001 Workshop on Models versus Exemplars in Computer Vision, December, 2001.
- [8] V. Blanz, T. Vetter: "A morphable model for the synthesis of 3D-faces". In Proc. of ACM SIGGRAPH, Los Angeles, 1999, pp187-194
- [9] S. Romdhani, V. Blanz, T. Vetter: "Face Identification by fitting a 3d Morphable Model Using Linear Shape and Texture Error Functions". In Proc. of European Conf. on Computer Vision, V4, 2002, pp3-19
- [10] V. Blanz, T. Vetter: "Face Recognition Based on Fitting a 3D Morphable Model". IEEE Pattern Analysis And Machine Intelligence, VOL.25, NO.9, September 2003
- [11] Yuxiao Hu, Dalong Jiang, Shuicheng Yan, Lei Zhang, Hongjiang zhang: "Automatic 3D Reconstruction for Face Recognition". IEEE International Conference on Automatic Face and Gesture Recognition (FGR' 04)
- [12] Dalong Jiang, Yuxiao Hu, Shuicheng Yan, Lei Zhang, Hongjiang Zhang, Wen Gao: "Efficient 3D reconstruction for face recognition". Pattern Recognition, Volume 38, Issue 6, June 2005, Pages 787-798
- [13] M. A. Oliver and R. Webster: "Kriging: a method of interpolation for geographical information system". Int. Journal on Geographical Information Systems, 4(3), 1990, pp313-332
- [14] Cartesia FACE SYSTEM <http://www.space-vision.jp/>