

時空間不変量と視点によらない動作認識

佐藤 淳[†]

Space-Time Invariants and Recognition of Motions from Arbitrary Viewpoints

Jun SATO[†]

あらまし カメラを用いて対象物を認識したりシーンの内容を理解する上で、それぞれの対象がもつ運動を識別することは非常に重要である。しかし、対象物の運動はカメラの視点が異なれば一般に異なる運動として観測される。本論文では、対象物の運動に関して視点によらない不変量が存在し、かつこれらが連続画像から計算可能であることを示す。まず、時間と空間がなす時空間を考え、実際の空間がなす時空間と画像に投影されたときにできる時空間との間での時空間投影に関し、その幾何学的な性質を明らかにする。更に時空間投影のもとでの不変量、すなわち時空間不変量が存在することを示す。実画像をもとに実際に時空間不変量を計算し、これが対象物の運動を識別する上で非常に有用であることを示す。更に、時空間不変量の画像ノイズに対する安定性についても報告する。

キーワード 時空間不変量, 幾何学的不変量, 動作認識, 時空間投影

1. ま え が き

対象物の運動に関する情報は、それぞれのシーンの内容を理解したり、その対象物自身を認識する上で非常に重要である [6]。例えば、手話やジェスチャのようにそれぞれの運動情報自身が意味をもっている場合もあるし、また人物を識別するときなどには、その人のもつ動きのくせなどが重要な役割を果たす場合もある。また医療においては、それぞれの病気や障害特有の症状の有無を判断することが求められるが、このような症状判断は必ずしも患者に現れる静的な特徴によるものとは限らず、時にはそれぞれの病気や障害特有の動作や行動をもとに病状を判断することが必要となる。コンピュータによる自動診断や診断支援システムにおいては、したがって、動作や運動の識別が非常に重要である。

これまでの研究から、時系列画像をもとに、人間の動的なアクションやジェスチャを認識したり、動きからそのシーンの文脈を理解するなどが可能であることが示された [3], [5]。これらの研究では統計的な特徴量をもとに運動を識別する方法がとられてきたが、これらの特徴量は視点が変化したり、カメラの内部パラ

メータが変化すると、一般に全く異なる値となるなどの問題があり、その応用には限界がある。

これに対し、任意の視点から対象物を認識する方法として、近年、幾何学的な不変量に基づく方法が注目されている [7]。多角形や多面体から曲線や曲面に至るまで様々な対象を識別するために、これまでに頂点などの特徴点をもとにした代数的不変量 [7] ~ [9], [13] や、曲線や曲面などの微分幾何学的な性質に基づく微分不変量 [4], [10] ~ [12] などが広く研究されてきた [1], [2]。しかし、これらの不変量はどれも対象物の形状や表面パターンなどの静的な特徴に関するものであり、対象物がもつ動的な特徴である動作や行動などを識別することはできなかった。

そこで本研究では時間と空間がなす時空間を考え、実際の時空間とこれを投影して得られる画像がなす時空間との間の関係を幾何学的にとらえることにより、時空間において幾何学的不変量が存在することを示す。時空間から時空間への投影は一般の変換とは異なり、各軸における変換が独立した異方性の投影である。このような時空間投影のもとでは一般に不変量は存在しないが、場合を限ることにより不変量を定義することが可能となる。本論文では、このような時空間投影のもとでの不変量、すなわち時空間不変量が存在する場合を示す。更に、このような時空間不変量を用いるこ

[†] 名古屋工業大学電気情報工学科, 名古屋市

Nagoya Institute of Technology, Nagoya-shi, 466-8555 Japan

とにより、対象物のもつ運動が視点によらず識別可能であることを示す。

以降の章では、まずはじめに各軸によって投影の性質が異なる異方性投影を考え、実際の時空間から画像がなす時空間への時空間投影がこのような異方性の投影であることを示す。次に、このような異方性の性質をもつ時空間投影に群の構造が存在することを示す。更に、このような群に関する不変量として時空間投影のもとでの運動に関する不変量の存在を示す。次に時空間不変量を一般化し、 N 点が M 時刻運動した場合の時空間不変量の分類について述べる。最後に、導出した時空間不変量を実際の連続画像に応用し、対象物のもつ運動が視点によらず識別可能であることを示す。

2. 異方性投影と時空間投影

図 1 (a1), (a2), (a3) は運動中の人物の連続画像の一部である。これらの連続画像において人物の手先をトラッキングし、2 次元画像に時間軸を加えた 3 次元の時空間においてその軌跡を表現したものを図 1 (a4) に示す。一方、同じ運動を異なるカメラ位置や姿勢で観測した結果が (b1) から (b4) である。投影された時空間における軌跡から人物の動作を識別したいのであるが、(a4) と (b4) に示すように、連続画像がなす時空間における軌跡は、カメラと人物との相対的な位置や姿勢が異なると全く違ったものとして観測される。本論文では、位置や姿勢やカメラの内部パラメータによらず同じ動作であれば同じ値となる時空間不変量をこれらの軌跡から計算する方法を導き出す。そこで、まず実際の空間がなす時空間から画像がなす時空間への投影の性質を調べる。

一般に空間から画像への投影は画像の x 軸に対しても y 軸に対しても同じ投影方法となっており、かつそれぞれの軸で独立ではない。例えば、平面物体が画像に透視投影されている場合には、3 次元空間中の平面 Π 上の点 $\mathbf{X} = [X, Y]^T$ から 2 次元画像面 π 上の点 $\mathbf{x} = [x, y]^T$ への投影は、次のように平面射影変換 P により記述できることはよく知られている。

$$\tilde{\mathbf{x}} \sim P\tilde{\mathbf{X}} \quad (1)$$

ここで、 $(\tilde{\cdot})$ は斉次座標を表し、また (\sim) は定数倍の不定性を除いて等しいことを表す。式 (1) では X 座標も Y 座標もともに射影変換によって投影されており、またそれぞれの変換は独立ではない。同様に弱透視投影を仮定する場合には、 X 座標の投影も Y 座標

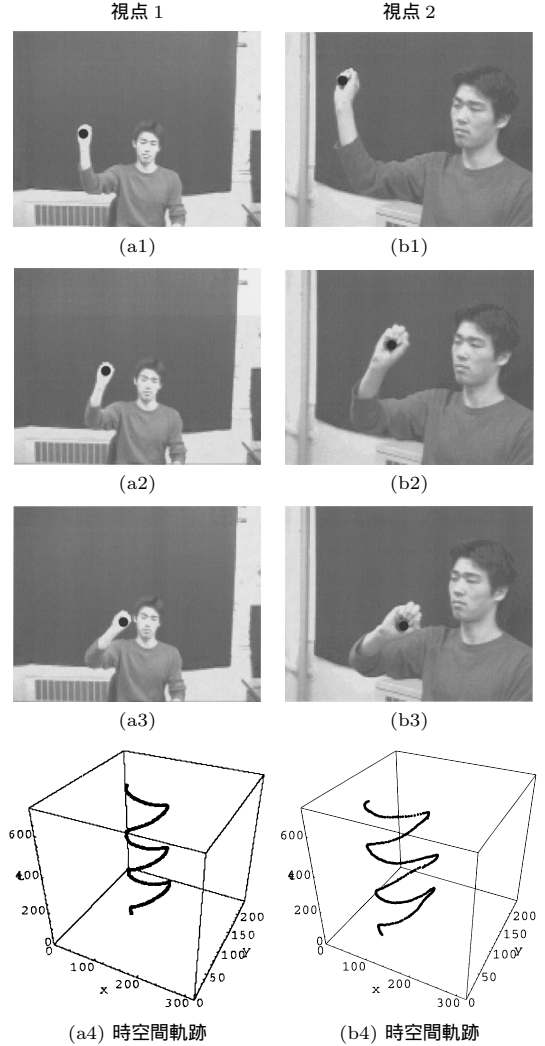


図 1 連続画像と時空間における軌跡
Fig. 1 Sequential images and trajectories in space-time.

の投影もともにアフィン変換によって記述できる。

しかし、すべての軸に対して同じ投影方法を採用するというのは投影法としては特殊なものであり、より一般的にはそれぞれの軸に対して異なる投影クラスを設定することが考えられる。例えば、 X 軸はアフィン変換で投影し、 Y 軸は射影変換で投影するといった投影方法をとることも考えられる。このように各軸が独立していてそれぞれの投影方法が異なるような投影を異方性投影と呼ぶことにする。異方性投影は時空間における投影を考える上で非常に重要である。

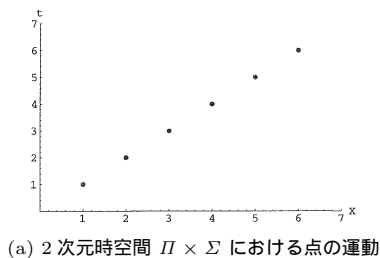
3. 時空間不変量の存在性

異方性投影のもとでは一般に不変量は存在しない。しかし、場合を限れば幾何学的な不変量を定義することが可能となる。今、 n 次元空間 Σ_1 に作用する変換 g_1 と、 m 次元空間 Σ_2 に作用する変換 g_2 があるとする。そしてこれらを組み合わせた変換 $g_1 \times g_2$ が $n+m$ 次元空間 $\Sigma_1 \times \Sigma_2$ に作用するとする。もしも

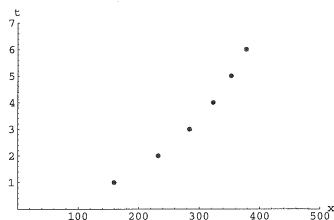
$$g_1 \times g_2 \in \mathcal{G}$$

なる変換群 \mathcal{G} が存在すれば、 \mathcal{G} のもとでの不変量を空間 $\Sigma_1 \times \Sigma_2$ 上において考えることができる。したがって、 g_1 を実空間 Σ_1 における投影とし、 g_2 を時間 Σ_2 に対する変換としたときに、もしも群 \mathcal{G} が存在すれば、時空間 $\Sigma_1 \times \Sigma_2$ における不変量が導出できると考えられる。

簡単のため、まず 1 次元の空間 Π と時間 Σ よりなる 2 次元の時空間 $\Pi \times \Sigma$ を考えることにする。図 2(a)



(a) 2次元時空間 $\Pi \times \Sigma$ における点の運動



(b) 投影された時空間 $\pi \times \sigma$ における点の運動

図 2 時空間投影。(a) は 1 次元空間と時間よりなる 2 次元の時空間における点の運動の例を示す。(b) はこの点の時空間画像への投影像である。もとの時空間において直線上に並ぶ点列は、投影された時空間においては一般に直線上には並ばず、両者の間に射影変換の関係はない

Fig. 2 Space-time projections. (a) shows the motions of a point in a 2D space-time. (b) shows projections of the motions in a spatio-temporal image. The linear motions in the space-time are not always projected to linear motions in the spatio-temporal image.

に示すように、1 次元空間 Π 中の点 $\mathbf{X} = [X]$ の運動は、2 次元時空間 $\Pi \times \Sigma$ における点 $\mathbf{W} = [X, t]^T$ の軌跡と考えることができる。

ここで時空間 $\Pi \times \Sigma$ を投影してできる時空間 $\pi \times \sigma$ を考える。1 次元空間 Π 中の点 \mathbf{X} の運動は時空間 $\pi \times \sigma$ 中では点 $\mathbf{w} = [x, t]^T$ の軌跡として観測される。すると、 $\Pi \times \Sigma$ 中の点 \mathbf{W} と $\pi \times \sigma$ 中の点 \mathbf{w} の間には 1 対 1 対応が存在する。ここで、 $\Pi \times \Sigma$ から $\pi \times \sigma$ への時空間投影 \mathcal{P} を考えることにしよう。

$$\mathcal{P}: \Pi \times \Sigma \rightarrow \pi \times \sigma \quad (2)$$

\mathcal{P} は空間 Π に対する 1 次元の射影変換 \mathcal{P}_1 と時間 Σ に対する 1 次元の恒等変換 \mathcal{P}_2 よりなる 2 次元の異方性投影である。 \mathcal{P}_1 も \mathcal{P}_2 もともに射影変換であるが、これらを組み合わせた $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ はもはや射影変換ではない。二つの独立な 1 次元射影変換を組み合わせた 2 次元の変換は一般に 2 次元射影変換とはならないからである。これは図 2 の時空間投影の例を見れば明らかである。この例では、2 次元時空間中の点 \mathbf{W} に対し、空間軸は $\begin{bmatrix} x \\ 1 \end{bmatrix} \sim \begin{bmatrix} 600 & 200 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix}$ なる射影変換で投影され、時間軸は恒等変換により投影されている。この図が示すように、 $\Pi \times \Sigma$ において直線上に並ぶ点は、 $\pi \times \sigma$ においては一般に直線上には投影されない。したがって、 $\Pi \times \Sigma$ から $\pi \times \sigma$ への射影変換は存在しない。

これを一般化して考えてみることにする。 n 次元空間 Π^n 中の点 \mathbf{x} に作用する n 次元射影変換 \mathcal{P}_1 と m 次元空間 Σ^m 中の点 \mathbf{y} に作用する m 次元射影変換 \mathcal{P}_2 をそれぞれ次式のように考える。

$$\lambda_1 \tilde{\mathbf{x}}' = \mathbf{P}_1 \tilde{\mathbf{x}} \quad (3)$$

$$\lambda_2 \tilde{\mathbf{y}}' = \mathbf{P}_2 \tilde{\mathbf{y}} \quad (4)$$

空間 $\Pi^n \times \Sigma^m$ 中の点 $\mathbf{z} = [\mathbf{x}^T, \mathbf{y}^T]^T$, $\mathbf{z}' = [\mathbf{x}'^T, \mathbf{y}'^T]^T$ に対して、もしも次式に示すような実数 λ_3 が存在すれば、 \mathcal{P}_1 と \mathcal{P}_2 を合成した $n+m$ 次元射影変換 \mathcal{P}_3 が存在する。

$$\lambda_3 \tilde{\mathbf{z}}' = \mathbf{P}_3 \tilde{\mathbf{z}} \quad (5)$$

しかし、このような実数 λ_3 は実際には存在しない。なぜならば、ある \mathbf{P}_1 と \mathbf{P}_2 に対して、式 (3) と式 (5) より λ_3 は \mathbf{x} のみに依存し、 \mathbf{y} には依存しない値であるはずなのに、式 (4) と式 (5) より λ_3 は \mathbf{y} のみに依存し \mathbf{x} には依存しない値でなければならないからで

ある．このような二つの条件をともに満たす実数 λ_3 を考えることはできない．

以上より，互いに独立な二つの射影変換 P_1 及び P_2 を合成した変換 $P = P_1 \times P_2$ は一般に射影変換とはならないことがわかる．時空間投影は， n 次元空間に作用する n 次元射影変換と，これとは独立な，時間に作用する 1 次元恒等変換とを合成したものである．恒等変換は射影変換の一つであるから，この場合の時空間投影は独立な二つの射影変換を合成したものである．よって，先の議論より時空間投影は一般に射影変換によって表すことはできないことがわかる．射影変換群は射影変換の閉じた集合であるから，時空間投影の集合は射影変換群には含まれない．したがって，このような時空間投影 P のもとでは，複比のような射影不変量は存在しない．

しかし，もしも空間軸の投影として弱透視投影が仮定できるならば，時空間投影のもとでの不変量を投影像から計算することが可能となる．弱透視投影や平行透視投影などはアフィン写像によって表すことができることから，この場合の時空間投影 P は n 次元の空間に作用する n 次元アフィン変換 P_1 と，時間に作用する 1 次元の恒等変換 P_2 よりなる $n+1$ 次元の異方性投影となる．恒等変換はアフィン変換の一種でもあるから，この場合の時空間投影は二つの独立なアフィン変換を合成したものと考えることができる．

そこで，一般の n 次元空間 Π^n に作用する n 次元アフィン変換 P_1 と，これとは独立な， m 次元空間 Σ^m に作用する m 次元アフィン変換 P_2 をそれぞれ次式のように考えることにする．

$$\begin{bmatrix} \mathbf{x}' \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{m}_1 \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} \mathbf{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_2 & \mathbf{m}_2 \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 1 \end{bmatrix} \quad (7)$$

ここで， \mathbf{M}_1 と \mathbf{M}_2 はそれぞれ $n \times n$ 行列と $m \times m$ 行列であり， \mathbf{m}_1 と \mathbf{m}_2 はそれぞれ n 次元と m 次元のベクトルである．また， $\mathbf{0}$ はゼロベクトルを表す．すると式 (6) 及び式 (7) より， P_1 と P_2 を合成した変換 $P = P_1 \times P_2$ は，次に示すような $n+m$ 次元アフィン変換となることがわかる．

$$\begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{O} & \mathbf{m}_1 \\ \mathbf{O} & \mathbf{M}_2 & \mathbf{m}_2 \\ \mathbf{0}^\top & \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ 1 \end{bmatrix} \quad (8)$$

ここで， \mathbf{O} はゼロ行列を表す．

以上より，二つの独立なアフィン変換を組み合わせた変換は，一般にアフィン変換となることがわかる．したがって空間軸の投影に弱透視投影を仮定した場合の時空間投影はアフィン変換となる．アフィン変換の集合はアフィン変換群をなし，アフィン変換群の作用のもとでは一直線上の 3 点の距離の比などのアフィン不変量が存在することがわかっている．よって，このような時空間投影の場合には，アフィン不変量として時空間不変量を定義することができる．

4. 時空間不変量の計算

以上より，空間軸の投影が弱透視投影である場合には，アフィン不変量として時空間不変量が計算可能であることがわかった．本章では，このような時空間不変量を代数的に計算する方法と幾何学的に計算する方法について述べる．

4.1 代数的計算法

まず，2 次元時空間 $\Pi \times \Sigma$ から 2 次元時空間 $\pi \times \sigma$ への時空間投影 $P = P_1 \times P_2$ を考える．空間軸の投影が弱透視投影である場合には，この時空間投影は 2 次元アフィン変換で表すことができ，通常の 2 次元アフィン変換のもとでの不変量として時空間不変量を定義することができる．したがって，投影された 2 次元時空間 $\pi \times \sigma$ 中に 4 点 \mathbf{w}_i ($i = 1, \dots, 4$) が存在する場合には，これらより計算した次のアフィン不変量 \mathcal{I} は時空間不変量となる．

$$\mathcal{I}(\mathbf{w}_i, \mathbf{w}_j, \mathbf{w}_k, \mathbf{w}_l) = \frac{|\tilde{\mathbf{w}}_i \quad \tilde{\mathbf{w}}_j \quad \tilde{\mathbf{w}}_l|}{|\tilde{\mathbf{w}}_i \quad \tilde{\mathbf{w}}_j \quad \tilde{\mathbf{w}}_k|} \quad (9)$$

ここで， $|\tilde{\mathbf{w}}_i \quad \tilde{\mathbf{w}}_j \quad \tilde{\mathbf{w}}_k|$ は三つのベクトル $\tilde{\mathbf{w}}_i, \tilde{\mathbf{w}}_j, \tilde{\mathbf{w}}_k$ を横に並べてできる 3×3 行列の行列式を表す．式 (9) において， i, \dots, l は $1, \dots, 4$ の順列より自由に選んでよい．

以上の解析はそのまま 3 次元の時空間へと拡張することができる．平面 Π^2 上の点 $\mathbf{X} = [X, Y]^\top$ の運動を考えよう．この点 \mathbf{X} が画像面 π^2 上で $\mathbf{x} = [x, y]^\top$ に投影されているとする．すると 3 次元時空間 $\Pi^2 \times \Sigma$ 中の点 $\mathbf{W} = [X, Y, t]^\top$ は画像がなす 3 次元の時空間 $\pi^2 \times \sigma$ において $\mathbf{w} = [x, y, t]^\top$ に投影される．弱透視投影が仮定できる場合には， $\pi^2 \times \sigma$ 中の 5 点 \mathbf{w}_i ($i = 1, \dots, 5$) より次に示す時空間不変量を計算することができる．

$$\mathcal{I}(\mathbf{w}_i, \mathbf{w}_j, \mathbf{w}_k, \mathbf{w}_l, \mathbf{w}_m) = \frac{|\tilde{\mathbf{w}}_i \quad \tilde{\mathbf{w}}_j \quad \tilde{\mathbf{w}}_k \quad \tilde{\mathbf{w}}_m|}{|\tilde{\mathbf{w}}_i \quad \tilde{\mathbf{w}}_j \quad \tilde{\mathbf{w}}_k \quad \tilde{\mathbf{w}}_l|} \quad (10)$$

ここで、 $|\tilde{w}_i \ \tilde{w}_j \ \tilde{w}_k \ \tilde{w}_l|$ は四つのベクトル $\tilde{w}_i, \tilde{w}_j, \tilde{w}_k, \tilde{w}_l$ を横に並べてできる 4×4 行列の行列式を表す。式 (10) において、 i, \dots, m は $1, \dots, 5$ の順序より自由に選んでよい。

4.2 幾何学的計算法

このような時空間不変量は次に示すように幾何学的に求めることもできる。図 3 に示すように 3 次元時空間中に 5 点 w_i ($i = 1, \dots, 5$) が存在するとする。すると、これら 5 点の内の任意の 3 点を通る平面 Γ と残りの 2 点を結ぶ直線との交点 w_6 を新たに定義することができる。例えば図に示すように w_1, w_2, w_3 を通る平面と w_4, w_5 を通る直線の交点 w_6 が得られる。このようにして平面上の 4 点 w_1, w_2, w_3, w_6 が求まると、今度はこれらのうちの 2 点同士を結ぶ直線の交点を得られ、一直線上に並ぶ 3 点が定義できる。例えば図 3 では、 w_1 と w_2 を結ぶ直線と、 w_3 と w_6 を結ぶ直線の交点 w_7 が得られ、一直線上の並ぶ 3 点 w_1, w_2, w_7 や w_3, w_6, w_7 が得られる。このようにして求めた一直線上の 3 点の距離の比、例えば $\frac{w_1 w_7}{w_1 w_2}$ などとはよく知られているようにアフィン不変量である [2]。このようにして、時空間中の 5 点から平面上の 4 点を求め、これらから直線上の 3 点を求めることにより時空間不変量を求めることができる。

4.3 関数的に独立な時空間不変量の数

式 (9) や式 (10) では、計算に用いる点 w_i, w_j, w_k, w_l, w_m の組合せ方により、関数的に独立な不変量が複数存在する。このような独立な時空間不変量の数について考える。

今、時空間中の任意の n 点の組 $p = \{w_1, w_2, \dots, w_n\}$ がなす空間 \mathcal{M} を考える。群 \mathcal{G} の任意の元 g による p の像の集合 $\mathcal{O} = \{g(p) | g \in \mathcal{G}\}$ を p を通る \mathcal{G} の軌道 (orbit) と呼ぶ。また、 p をそれ自身に写像す

ような変換がなす群 $\mathcal{G}_p = \{g \in \mathcal{G} | g(p) = p\}$ を、 \mathcal{G} の p における固定群と呼ぶ。群 \mathcal{G} の次元を $\dim \mathcal{G}$ 、固定群 \mathcal{G}_p の次元を $\dim \mathcal{G}_p$ とすると、軌道 \mathcal{O} の次元 $\dim \mathcal{O}$ は次のように表すことができる。

$$\dim \mathcal{O} = \dim \mathcal{G} - \dim \mathcal{G}_p \quad (11)$$

空間 \mathcal{M} の次元を $\dim \mathcal{M}$ とすると、関数的に独立な不変量の数 m は一般に

$$m = \dim \mathcal{M} - \dim \mathcal{O} \quad (12)$$

と表せる [1], [7]。

2 次元時空間における時空間不変量の場合には、式 (9) に示すとおり 2 次元平面上の 4 点より不変量を計算するため \mathcal{M} の次元は 8 であり、また軌道 \mathcal{O} の次元は 2 次元アフィン変換群の次元と等しく 6 である。したがってこの場合には関数的に独立な不変量が $8 - 6 = 2$ 個存在する。また、3 次元時空間における時空間不変量の場合には、式 (10) に示すように 3 次元空間中の 5 点より不変量を計算するため \mathcal{M} の次元は 15 であり、軌道 \mathcal{O} の次元は 3 次元アフィン変換群の次元と等しく 12 である。したがってこの場合には関数的に独立な不変量が $15 - 12 = 3$ 個存在することがわかる。

4.4 時間軸サンプリングの同一性

以上述べたとおり、時空間という拡張された空間において不変量を考えることにより、これまで形状などの静的な特徴の認識に限定されていた不変量を、運動や行動などの動的特徴の認識に拡張して応用することが可能となる。しかし、時間という新たな次元を入れたことにより、不変量計算のための特徴点の取得時に、時間軸方向のサンプリングの同期を保つことが新たに必要となる。

提案した時空間不変量は、平面運動の場合には時空間中の 5 点から計算されるが、これらの点は複数の視点において同じタイミングでサンプルされた点であることが必要である。例えば、第 1 の視点では 1 秒間隔でサンプリングし、第 2 の視点では 5 秒間隔でサンプリングした場合には、同一の運動を観測したにもかかわらず、これら二つの視点で計算した時空間不変量は一般に一致しないことに注意が必要である。

また、同一間隔でサンプリングしてもサンプリングの開始時刻が異なればやはり不変量は一致しない。しかし、時間軸方向のサンプリングが十分に密であれば、同一の不変量をもつ点を時間軸方向に探索していくことで、同一の運動をもつ部分を探すことができる。こ

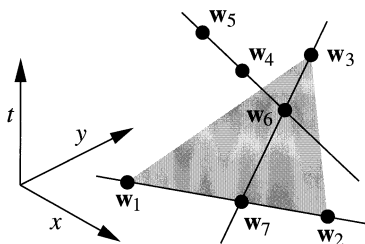


図 3 時空間不変量の幾何学的計算法

Fig. 3 Geometric method for computing space-time invariants.

のような探索は、時空間不変量に限らず、ある特徴量を用いて同一物を探索する場合に必ず必要となる処理である。

5. 時空間不変量の分類

以上述べた時空間不変量は、平面上を運動する単一点の 5 時刻の画像から計算してもよいし、平面上に存在する複数の点がそれと同一の平面上を運動する場合には、これらの複数点を用いることにより、より少ない時刻の画像から時空間不変量を計算することもできる。点の数と必要な時刻数はトレードオフの関係にあり、以下に述べるように、点の数と時刻数によって時空間不変量は運動情報のみに関する不変量となったり、運動と形状の双方に関する不変量となったりする。

時空間不変量を計算するには、時空間中に 5 点が存在すればよい。したがって、以下の条件が成り立つ場合には時空間不変量が求まる。

$$[\text{空間中の点の数}] \times [\text{時刻数}] \geq 5 \quad (13)$$

式 (13) をもとに、空間中の点の数と不変量を計算するのに必要な時刻数をまとめたものを表 1 に示す。

時空間不変量は点の数と時刻数に応じて、純粋に運動のみに関する不変量となったり、運動と形状の両方に関する不変量となったりする。運動情報のみに関する不変量とは、形状に依存せず運動が等しいときに同じ値となる不変量であり、運動と形状に関する不変量とは、対象物の運動と形状の双方が等しいときに同じ値となり、運動と形状のうちどちらか一方でも異なる場合には一般に異なる値となる不変量である。1 点 5 時刻から求めた時空間不変量は純粋に運動情報のみに関する不変量となり、2 点 3 時刻や 3 点 2 時刻から求まる時空間不変量は、運動と形状の両方に関する不変量となる。また 5 点 1 時刻から計算した不変量は、これまでによく知られているように純粋に形状のみに関する不変量となる (表 1 参照)。

表 1 点の数と時刻数と不変量のタイプ

Table 1 The relationship between the number of points, the number of time and the type of space-time invariants.

点数	時刻	不変量のタイプ
1	5	運動のみ
2	3	運動 + 形状
3	2	運動 + 形状
4	2	運動 + 形状
5	1	形状のみ

このように、時空間という拡張された空間において不変量を考えることにより、これまで形状のみに限定されていた不変量に対し、運動に関する不変量や、運動と形状に関する不変量などを新たに定義することが可能となる。

6. 時空間不変量曲線

以上のとおり、平面運動に関する時空間不変量は時空間中に最低 5 点が存在すれば求めることができる。しかし、実際に運動を観測すると時空間中には多くの点が得られる。これらの点を用いることにより、不変量を単一の数値としてではなく、多次元ベクトルや曲線として表現することができる。

今、ある時間間隔 t_1, t_2, t_3, t_4 を適当に決める。これらの数値は自由に選んでよいが、ある固定した値とする。ある時刻 t における時空間中の点を \mathbf{w}_t で表すことにすると、時刻 t における時空間不変量 $\mathcal{I}(t)$ は $\mathbf{w}_t, \mathbf{w}_{t+t_1}, \mathbf{w}_{t+t_2}, \mathbf{w}_{t+t_3}, \mathbf{w}_{t+t_4}$ の 5 点から次のように式 (10) より計算することができる。

$$\mathcal{I}(t) = \mathcal{I}(\mathbf{w}_t, \mathbf{w}_{t+t_1}, \mathbf{w}_{t+t_2}, \mathbf{w}_{t+t_3}, \mathbf{w}_{t+t_4}) \quad (14)$$

このように各時刻 t において時空間不変量 \mathcal{I} を計算し、これをグラフ化することにより、不変量の曲線を得ることができる。不変量曲線は同一の運動に関してはカメラの位置や姿勢によらず不変となる。このように時空間不変量を多次元化あるいは曲線化することにより、単一の数値の場合に比べてより識別能力が向上することが考えられる。

7. 実験結果

次に時空間不変量に関する実験結果を示す。

7.1 単一点運動からの時空間不変量の計算

まず、工業製品の機械的な運動から時空間不変量を計算した結果を示す。図 4(a), (b) はメトロノームの振り子の運動を二つの異なる視点に置いたカメラで撮影したときの画像における軌跡である。この画像上の軌跡を時空間において表現したものが図 4(c), (d) である。二つの視点におけるそれぞれの時空間軌跡から提案した時空間不変量を計算した結果を図 4(e) に示す。(e) に見るように、同じ運動に関する時空間不変量は視点が異なってもほぼ一致している。図 5 は先とは異なる時計の針の運動の例である。同様に二つの異なる視点において計算した時空間不変量は非常によく一致している。また、図 4(e) 及び図 5(e) から明らか

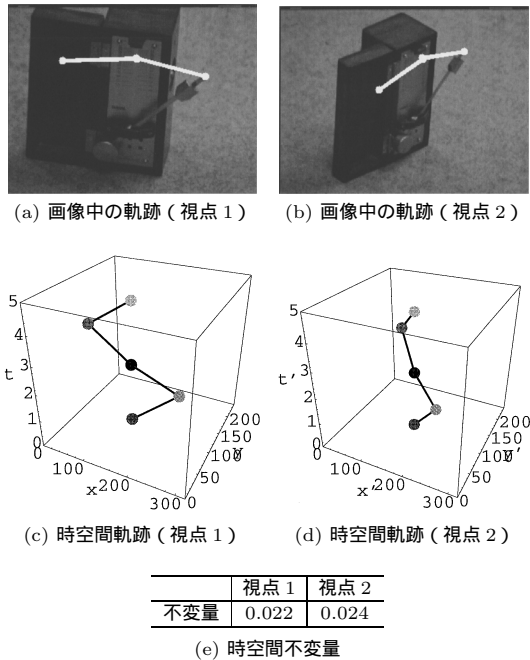


図 4 時空間不変量の計測結果 (メトロノーム)

Fig. 4 The space-time invariants computed from mechanical motions (metronome).

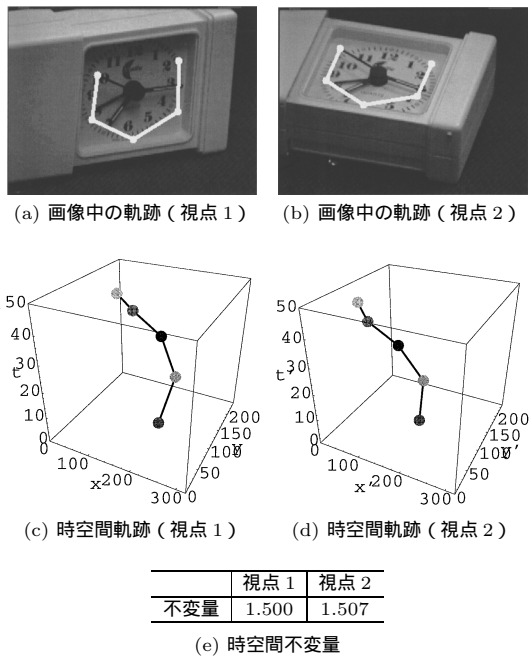


図 5 時空間不変量の計測結果 (時計)

Fig. 5 The space-time invariants computed from mechanical motions (clock).

のように、二つの運動の時空間不変量は大きく異なっている。

7.2 時空間不変量曲線の計算

次に長時間の運動から時空間不変量曲線を計算した結果を示す。図 6 (a) 及び (b) は、十字を描く手の運動を二つの異なる視点において観測した結果である。白線は関連トラックにより得られた運動軌跡である。視点が異なれば見かけの運動も大きく異なる。それぞれの視点における連続画像から時空間における運動軌跡を求めた結果を図 6 (c) 及び (d) に示す。この手の運動を平面運動とみなし、弱透視投影を仮定して式 (10)

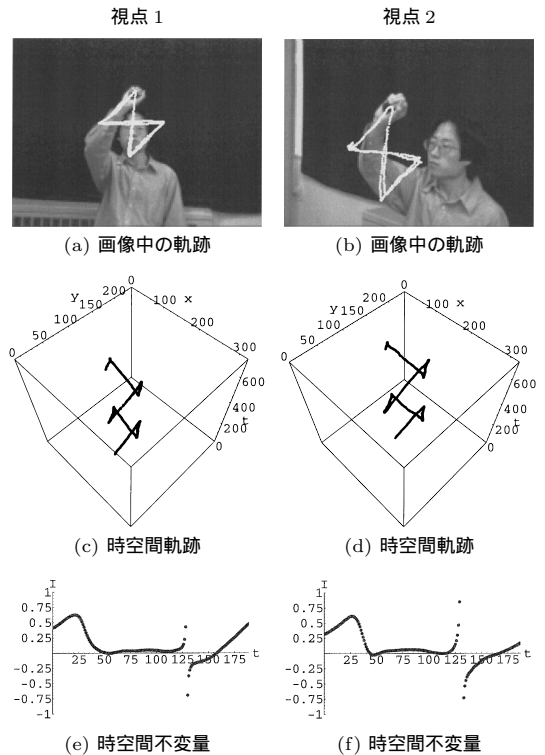


図 6 時空間不変量の計測結果。(a) 及び (b) は二つの異なる視点において十字を描く手の運動を観測した例である。図中の白線は関連トラックにより運動軌跡を抽出した結果である。(c) 及び (d) は二つの視点における画像時空間中での運動軌跡である。(e) 及び (f) それぞれから計算した時空間不変量を (e) 及び (f) に示す。画像中の見かけの運動が異なるにもかかわらず、時空間不変量はほぼ一致していることがわかる。

Fig. 6 The space-time invariants computed from human motions. As shown in (e) and (f), space-time invariants computed from two different viewpoints are almost identical.

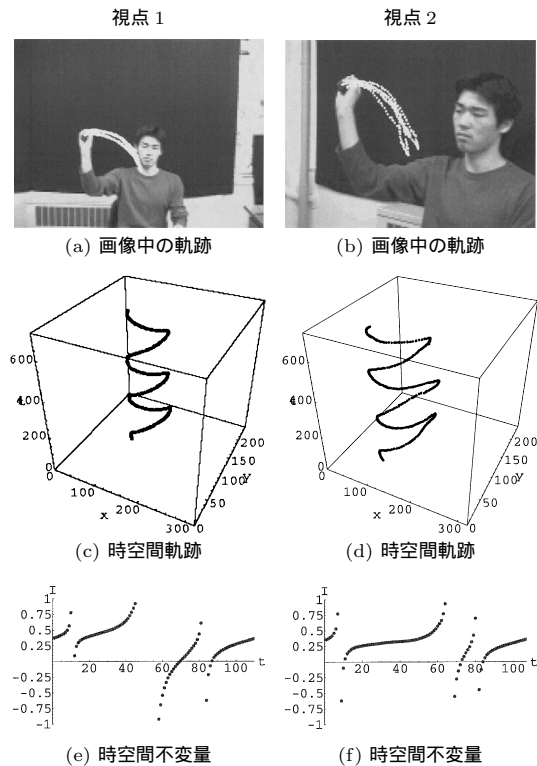


図 7 時空間不変量の計測結果．(a) 及び (b) は手を左右に振る運動を二つの異なる視点において観測した例である．以下，図 6 と同様に，時空間における軌跡を (c) 及び (d) に，また時空間不変量の計算結果を (e) 及び (f) に示す

Fig. 7 The space-time invariants computed from human motions.

よりそれぞれの運動軌跡から時空間不変量曲線を計算した結果を図 6 (e), (f) に示す．(e) 及び (f) に見るように，二つの異なる視点における時空間不変量はほぼ一致している．単一の不変量値に比べ，このように多数の不変量を計算して曲線化することにより，運動識別能力がより向上すると考えられる．

図 7 (a) 及び (b) は手を左右に振った場合の例である．同様に時空間不変量の曲線を計算した結果を図 7 (e), (f) に示す．これらのグラフより明かなように，十字運動の場合と左右運動の場合の不変量曲線は大きく異なっており，時空間不変量により二つの異なる運動が識別可能であることがわかる．

7.3 複数点運動からの時空間不変量の計算
次に，複数の点が運動したときに，運動と形状に關する時空間不変量を計算した結果を示す．図 8 (a) は

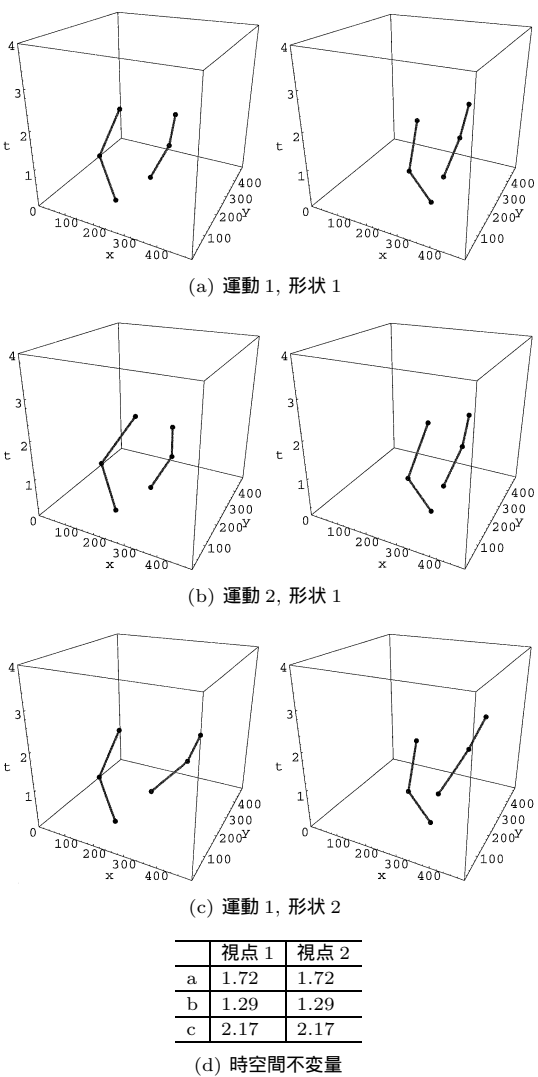


図 8 2 点 3 時刻による時空間不変量
Fig. 8 The space-time invariants computed from two points in three sequential time.

二つの点が 3 時刻の間回転運動したときの二つの異なる視点における画像中の軌跡である．(b) は (a) と同じ 2 点が (a) とは異なる運動を行った場合であり，(c) は (a) とは異なる 2 点が (a) と同じ運動を行った場合である．それぞれにおいて 2 点 3 時刻の軌跡から不変量を計算した結果を (d) に示す．(d) より，この場合の時空間不変量は，運動と形状のうち一方でも異なれば，異なる値となることがわかる．また，運動と形状が等しい場合には，視点が異なっても同じ値となって

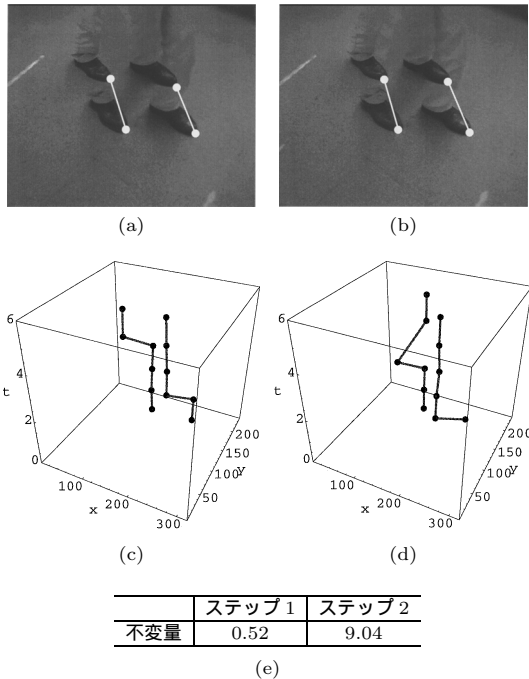


図 9 二つの異なるダンスステップから計算した時空間不変量

Fig. 9 The space-time invariants computed from two different dance steps.

いる．このように，複数点の複数時刻の運動からは，時空間不変量により視点によらず運動と形状を同時に識別することができる．

図 9 (a), (b) は二つの異なるダンスステップの連続画像を 6 時刻分重ね合わせて表示したものであり，左右の足の爪先の軌跡が白線で示してある．これら二つの運動は，画像に投影された軌跡は全く同じであるが，足を出すタイミングがそれぞれで異なっているため，実際には異なる運動である．図 9 (c), (d) はこれら二つのステップの時空間軌跡である．これらから時空間不変量を計算した結果を (e) に示す．二つの時空間不変量は大きく異なっており，二つのステップが不変量を用いて確実に識別できることがわかる．

7.4 時空間不変量の安定性

次に時空間不変量の耐ノイズ性を評価した結果を示す．図 8 の (a), (b), (c) 三つの画像軌跡に対し標準偏差 1 ピクセルのランダムノイズを加えて時空間不変量を計測した結果を図 10 に示す．横軸は図 8 の (a), (b), (c) の三つの運動を表し，縦軸は時空間不変量の値である．不変量の平均は黒丸で，また 3σ の不確定

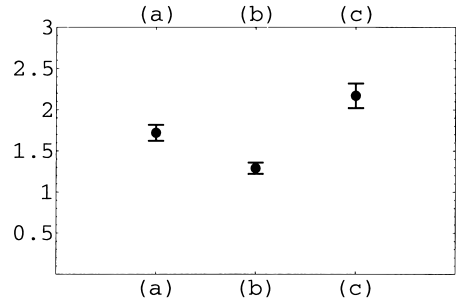


図 10 時空間不変量の耐ノイズ性
Fig. 10 Stability of space-time invariants.

領域は線で示す．不確定領域に重なりがないことから，時空間不変量は十分に運動識別能力があることがわかる．しかし，人間のジェスチャなどを識別しようとする場合には，もともとの運動の再現性があまりよくないことから，不変量の不確定領域はより大きくなることが考えられる．

8. 同一性と不変量

不変量を用いて物体認識や識別を行う場合には，視点に依存しないという大きな長所がある半面，不変量が一一致するということが我々人間が同一であると感じる感覚とが必ずしも一致しないという問題も存在する．

一般に人間は形状が等しいということをユークリッド変換群における不変量としてとらえている．これに対し，アフィン不変量や射影不変量を用いる場合には，アフィンの，あるいは射影的な同一性を見ていることになる [2]．ユークリッド変換群は射影変換群やアフィン変換群の部分群であるから，ユークリッド的に等しいものは，アフィンのにも射影的にも必ず等しいことが保証されるが，逆にアフィンの，あるいは射影的に等しいものであっても，必ずしもユークリッド的に等しいとは限らない．アフィン不変量や射影不変量を用いて同一性を議論する場合には，我々人間が定義するユークリッド的な同一性とこのような多少の違いがあることを意識しなければならない．

したがって，厳密に考えた場合，本論文で示した時空間不変量においては，同一の運動は必ず等しい不変量をもつことが保証されているが，逆に不変量が等しいからといっても必ずしも運動が同一であるとはいえないという問題が存在する．しかし現実の認識問題では多くの場合，非常に広範な運動の可能性の中からある特定の運動を見つけるという作業になる．このよう

な場合、アフィン不変量が一致するという事は、かなりの確度でユークリッド的にも一致していることを示しており、同一の運動であると考えても実用上問題ない場合が多い。系全体がもつ自由度に比べて、アフィン変換とユークリッド変換の差によって生ずる自由度の方がはるかに小さいからである。

9. む す び

本論文では時空間における不変量をもとにシーン中の対象物の運動を識別する新しい方法を提案した。

まず、時空間投影では時間軸の投影という異質な投影が加わることにより、通常の等質な N 次元投影とは異なる異方性の投影となることを示した。このような時空間投影のもとでは射影不変量は存在しないが、弱投影投影を仮定することにより、アフィン不変量として時空間不変量が求まることを明らかにした。更に、空間中の点の数と時刻数によって、時空間不変量の性質が異なることも示した。

実画像を使った実験により、時空間不変量が実際に計算可能であることを確認した。更にシミュレーションにより、提案した時空間不変量の画像ノイズに対する安定性を示した。

本論文では平面運動を扱ったが、一般の3次元空間中の運動に関する時空間不変量を考える場合には、4次元の時空間から画像がなす3次元の時空間への投影を扱うことになり次元の縮退が生ずる。このような3次元空間中の運動への拡張が今後の課題である。

謝辞 本研究に関し多くの御助言を頂いた名古屋工業大学の佐藤幸男教授に感謝致します。なお本研究は、栢森情報科学振興財団研究助成 (K10 研 III-43) 及び日本学術振興会科学研究費補助金 (奨励研究 A-11780263) を受けて行った。

文 献

- [1] 杉本晃宏, “幾何学的不変量とその応用,” コンピュータビジョン—技術評論と将来展望, 松山隆司, 久野義徳, 井宮 淳 (編), pp.97–122, 新技術コミュニケーションズ, 1998.
- [2] 佐藤 淳, コンピュータビジョン—視覚の幾何学, コロナ社, 1999.
- [3] A.F. Bobick and A.D. Wilson, “A state-based technique for the summarization and recognition of gesture,” Proc. 5th International Conference on Computer Vision, pp.382–388, Cambridge, USA, 1995.
- [4] D. Cyganski, J.A. Orr, T.A. Cott, and R.J. Dodson, “An affine transformation invariant curvature function,” Proc. 1st International Conference on Computer Vision, pp.496–500, London, 1987.

- [5] T. Darrell and A. Pentland, “Space-time gestures,” Proc. Conference on Computer Vision and Pattern Recognition, pp.335–340, New York, 1993.
 - [6] J.J. Gibson, The Ecological Approach to Visual Perception, Houghton Mifflin, 1979.
 - [7] J.L. Mundy and A. Zisserman, Geometric Invariance in Computer Vision, MIT Press, Cambridge, USA, 1992.
 - [8] L. Quan, “Invariants of six points and projective reconstruction from three uncalibrated images,” IEEE Trans. Pattern Anal. & Mach. Intell., vol.17, no.1, pp.34–46, 1995.
 - [9] C.A. Rothwell, A. Zisserman, D.A. Forsyth, and J.L. Mundy, “Planar object recognition using projective shape representation,” International Journal of Computer Vision, vol.16, pp.57–99, 1995.
 - [10] J. Sato and R. Cipolla, “Quasi-invariant parameterisations and matching of curves in images,” International Journal of Computer Vision, vol.28, no.2, pp.117–136, 1998.
 - [11] L.J. Van Gool, T. Moons, E. Pauwels, and A. Oosterlinck, “Semi-differential invariants,” in Geometric Invariance in Computer Vision, ed. J.L. Mundy and A. Zisserman, pp.157–192, MIT Press, 1992.
 - [12] I. Weiss, “Geometric invariants and object recognition,” International Journal of Computer Vision, vol.10, no.3, pp.207–231, 1993.
 - [13] A. Zisserman, D. Forsyth, J. Mundy, C. Rothwell, J. Liu, and N. Pillow, “3D object recognition using invariance,” Artif. Intell., vol.78, pp.239–288, 1995.
- (平成12年9月25日受付, 13年1月31日再受付)



佐藤 淳 (正員)

昭59名工大・工卒・平8ケンブリッジ大学院博士課程了。同年ケンブリッジ大・工・助手。平10名工大・工・助教授。この間に、ATR人間情報通信研究所客員研究員など。コンピュータビジョン、幾何学的不変量、視覚誘導、視覚的ユーザインタフェースの研究に従事。博士 (PhD)。著書に「コンピュータビジョン—視覚の幾何学—」(コロナ社) など。BMVC'94 最優秀科学論文賞, BMVC'97 最優秀科学論文賞, 第5回画像センシングシンポジウム論文賞等各受賞。情報処理学会会員。