コンピュータビジョン　93－3
（1995．3．23）

CVCV-WG 特別報告*: コンピュータビジョンにおける技術評論と将来展望 (II)
— ビジョンにおける不変量とその応用 —

杉本 晃宏
ATR 人間情報通信研究所

ある幾何学的対象がある変換に従うとき，その対象を特徴づけるパラメタによって定義され，対象が従う変換に影響されない値をとる関数が不変量である．ビジョンの枠組では，対象が画像面に投影される変換 (透視変換) に対して，不変量を求めることに関心が寄せられる．こうした不変量は，物体認識をはじめとするコンピュータビジョンにおける多くの課題に有効な手段を提供するからである．本稿では，不変量に関する最近の研究成果を概観し，コンピュータビジョンにおける不変量の研究について今後の方向性を探る．まず，これまでに得られている不変量を紹介し，その幾何的導出，および，代数的導出を解説する．そして，不変量の応用例を示し，そこにおける特長と問題点について述べる．さらに，不変量を利用するに際して，その限界と今後の展望について述べるとともに，CVCV-WG で行われた議論を掲載する．

CVCV-WG Special Report: Technical Review and View in Computer Vision (II)
— Invariants within the Context of Vision and Their Applications —

Akihiro SUGIMOTO
ATR Human Information Processing Research Laboratories
Soraku-gun, Kyoto 619-02, Japan
e-mail: sugimoto@hip.atr.co.jp

An invariant of a geometric configuration is a function of the configuration whose value is unchanged by a particular transformation. Within the context of vision, we are interested in finding invariants under the perspective projection of an object into an image; such invariants provide many machine vision tasks, such as object recognition, with a powerful platform. This paper presents recent results on invariants in computer vision and searches for the direction we should take in the future research of invariants. We first describe invariants obtained so far from both geometric and algebraic aspects. We then give some applications of invariants to machine vision tasks and summarize the advantages and the problems in using invariants. The limitation and the perspective of usage of invariants are also presented. Finally, discussion by the CVCV-WG members on the breakthrough to the next stage where invariants are much more useful is given.

---

*情報処理学会コンピュータビジョン研究会では，CV 技術評論・将来展望ワーキンググループ (略称 CVCV-WG, Current and Vision of Computer Vision, 幹事: 井宮 淳 (千葉大), 久野 義徳 (阪大)) の活動を開始した．この WG では，1980 年代からの CV 研究の総括を行い，「何ができて何ができないのか，またその理由は何か」を具体的アルゴリズムのレベルから詳細に見直すことによって，今後の vision 研究の方向を見出すことを目的とする．CV の新潮流を探るために若手数名が批判的なポジションペーパー ＋ サーベイを作成し，それを基に WG で議論を行い，今後の方向性を探る予定である．今回の報告はその第 2 回目にあたる．なお，本活動についてのご意見，ご質問は，cvcv@top.it.okayama-u.ac.jp まで．

1　はじめに

“投影を通して得られる2次元投影像から，3次元的に存在する対象に対して何がいえるか”はコンピュータビジョンにおける中心課題の一つである．投影像を用いて3次元物体を考察しようとしたときに生じる本質的な問題は，1)3次元から2次元への投影によって情報が欠落すること，2)視点(投影の中心)に依存して3次元物体の見え方が変わること，である．物体は3次元情報をもつが投影像は2次元情報しかもたないため，1枚の投影像を用いる限り，1)の問題を克服することはできない．そこで，なめらかさなどの適当な仮定を設けて正則化手法[34]を利用し，あるいは複数の投影像を扱うことで欠落した情報を補い，この問題に対処している．一方，上記2)により同じ物体であってもその投影像は無数に存在し，3次元物体を知るにはそれらをどううまく取り扱うかがカギとなる．そこで，一つの物体に対して数枚の投影像をテンプレートとして用意し，そしてそれらを組み合わせてその物体の他の投影像を表現し，投影像の取り扱いに対処する試み[52],[55]もあるが，投影像から視点に依存しない情報を抽出し，これを投影像の取り扱いに利用しようという気運が高まってきている[10],[31],[57]．視点に依存しない情報は，物体と投影像を関係づける本質的な量の一つであると考えられるからである．また，視点に依存しない情報を利用すれば，物体を蓄積したデータベースに高速にアクセスできる[13],[44]からである．

ある幾何学的対象がある変換に従うとき，その対象を特徴づけるパラメタによって定義され，対象が従う変換に影響されない値をとる関数が不変量である．例えば，2点間の距離は，ユークリッド変換(回転と並進)のもとで不変であるが，相似変換(回転，並進と一様なスケーリング)のもとでは不変でない．これに対し，角度は，ユークリッド変換，相似変換，いずれのもとでも不変である．ビジョンの枠組では画像面への投影は透視変換によって行われるため，視点に依存しない情報はこの透視変換下での不変量となる．したがって，この透視変換下での不変量を求めることがビジョンの分野における不変量の研究となる．

一方，Klein [24]によって“幾何学は，与えられた変換群の下での，空間の不変な性質を研究することである”という思想が掲げられ，これをうけて，不変量に関する研究は19世紀後半に精力的におこなわれた．しかし，ここで研究された不変量[17],[33]には，“投影による情報の欠落”の問題が介在していない．すなわち，ここでは3次元情報そのものを取り扱う枠組で不変量が研究され，投影を通して得られる2次元情報を取り扱う枠組で不変量が論じられていない．上に述べたようにビジョンの枠組では，3次元情報を直接扱うことができず，投影されて得られる2次元情報を扱わなければならない．Klein [24]の卓見にもとづいた不変量の研究とビジョンにおける不変量の研究とはこの点で大きく異なる[1]．

本稿では，コンピュータビジョンにおける不変量に関する最近の研究成果を概観する．また，不変量をコンピュータビジョンの課題に応用した例を示し，そこにおける特長と問題点について述べる．さらに，ビジョンにおける不変量の研究に関して，克服すべき課題について述べ，今後の方向性を探る．以下，第2節において，これまでに得られている不変量を解説する．まず，不変量を数学的に定義し，与えられた幾何学的対象とそれに対して存在しうる不変量の数の関係について述べる．次に，最近得られた主な不変量を紹介し，その幾何的導出，および，代数的導出を解説する．ここでは，1枚の投影像から得られる不変量と複数枚の投影像から得られる不変量とに分けて述べる．第3節では，コンピュータビジョンの課題への不変量の応用例として，不変量を利用した認識システムをとりあげ，そこにおける特長と問題点を指摘する．第4節では，不変量を利用するに際しての限界と，不変量をより利用可能(実用的)なものにするための今後の課題について述べる．CVCV-WGのメンバーによって行われた議論の要約を第5節に掲載する．

2　不変量

本節では，ビジョンの枠組での不変量を数学的に定義し，最近導出された主な不変量を1枚の投影像から得られるものと複数枚の投影像から得られるものに分類して紹介する．そして，それらの幾何的導出，および，代数的導出を解説する．さらに，微分不変量と準不変量について簡単に触れる．なお本節では，投影像が複数ある場合，1)投影像間の特徴(点や直線など)の対応は既知であること，2)不変量の導出に要する特徴に関してはオクルージョンが生じていないこと，を仮定する．

2.1　不変量とは

ビジョン研究での関心は，投影モデルである透視変換のもとでの不変量を求めることである．不変量は対象の，与えられた変換群に依存しない性質であるが，透視変換は群をなさないので，我々が関心を寄せる不変量はKlein [24]の意

[1] ビジョンにおける不変量の研究は，過去に得られた不変量をビジョンの枠組で利用可能なものに読き直すこと，あるいは，過去に得られた不変量を実用可能な形に整備することでもある．

味での不変量の枠組に属さない．そこで，ビジョンにおける不変量の研究を行うためには，ビジョンの枠組に沿って不変量を数学的に定義しなおさなければならない．すなわち，“投影による情報の欠落” を考慮した枠組で不変量を定義しなければならない．

対象が存在する 3 次元ユークリッド空間を 3 次元射影空間に，画像面を射影平面に埋め込むことによって，透視変換を 3 次元から 2 次元への射影変換として表すことができる．ユークリッド空間内で議論する限り透視変換は非線形な変換であるが，ユークリッド空間を射影空間に埋め込むことによって，3 次元空間と画像面を線形な関係で結ぶことができる．さらに，射影空間を考えることにより，キャリブレーション (カメラの内部パラメタと外部パラメタの決定) が行われていることを前提としないで議論できるという利点がある．ユークリッド空間内で透視変換を正射影や一様なスケーリングをともなう正射影で近似したり，ユークリッド空間をアフィン空間に埋め込み，透視変換をアフィン変換で近似したりして，3 次元空間と画像面を線形な関係で結ぶ試みもあるが，本稿では透視変換を近似しないという立場をとる．

n を自然数として，$\mathcal{P}^n$ をある体 (本稿では実数体 $\mathbf{R}$) 上の n 次元射影空間とする．3 次元空間内でのユークリッド座標が $\tilde{\boldsymbol{x}} = (x,y,z)^{\mathrm{T}}$ である[2]点の斉次座標が $\boldsymbol{x} = (1,x,y,z)^{\mathrm{T}}$ となるように 3 次元ユークリッド空間を $\mathcal{P}^3$ に埋め込む．同様に，画像面上でのユークリッド座標が $\tilde{\boldsymbol{X}} = (X,Y)^{\mathrm{T}}$ である点の斉次座標が $\boldsymbol{X} = (1,X,Y)^{\mathrm{T}}$ となるように画像面を $\mathcal{P}^2$ に埋め込む．以下では，特に断わらない限り，座標とは斉次座標を意味する．また，射影幾何学に関しては，[7] や [10] などを参照されたい．座標 $\boldsymbol{x}$ にある点の透視変換による像の座標を $\boldsymbol{X}$ とすると，

$$\boldsymbol{X} \;=\; f_P(\boldsymbol{x}) \tag{2.1}$$

が成り立つ．ここに f_P は，$\mathrm{rank}P = 3$ なる 3×4 の行列 P によって表現される射影変換である．なお，透視変換を実現するピンホールカメラのパラメタに関する情報はすべて，P の要素に含まれる．すなわち，式 (2.1) に基づく議論ではキャリブレーションを必要としない．また，P のカーネルが視点 (投影の中心) の座標に一致する．$\mathcal{P}^3$ から $\mathcal{P}^2$ への射影変換を表す 3×4 の行列 (ただし，そのランクは 3) 全体の集合を $\mathcal{C}$ とおくと，$\mathcal{C}$ の要素は，あるカメラで，ある位置から 3 次元物体を画像面に投影する変換 (透視変換) に対応する[3]．ビジョンにおける不変量とは，観測している 3 次元物体が $\mathcal{C}$ の任意の要素で変換を受けても，それに依存しない値を有する関数である．しかもその関数は，その物体の投影像から得られる量 (すなわち，$\mathcal{C}$ の要素による変換後の座標) を用いて定義されなければならない．

$\mathcal{P}^3$ の部分集合[4]を S とし，$P \in \mathcal{C}$ (ただし P のカーネルは S に含まれない) に対して $f_P(S) := \bigcup_{\boldsymbol{x}\in S}\{f_P(\boldsymbol{x})\}$ とする．ビジョンの枠組での S に対する不変量は次のように定義され，我々の目的は，“与えられた S と N に対して，関数 I を求めること” である．なお，N は不変量を導出するのに要する投影像の数に対応する．

定義 2.1 $\mathcal{P}^3$ の部分集合 S が与えられているとする．そして，自然数 N に対して，$P_i \in \mathcal{C}$ (P_i のカーネルは S に含まれない; $i = 1,2,\ldots,N$) が与えられたとする．このとき，ビジョンの枠組での S に対する不変量は，P_i ($i = 1,2,\ldots,N$) に依らない値をとる関数 $I:\ f_{P_1}(S) \times f_{P_2}(S) \times \cdots \times f_{P_N}(S) \longrightarrow \mathbf{R}$ である． □

注意 2.1 集合 S と S に作用する変換群 $\mathcal{G}$ が与えられているとする．このとき，$\mathcal{G}$ のもとでの S に対する Klein [24] の意味での不変量は，$\mathcal{G}$ の各軌道空間上で一定の値をとる関数 $Inv:\ S \rightarrow \mathbf{R}$ である．ここに $\mathcal{G}$ の軌道空間とは，“S の 2 点が $\mathcal{G}$ の適当な要素によって他に移りあえる (同じ軌道上にある) とき，この 2 点は同値である” という同値関係による S の商空間である．また，このとき $\boldsymbol{x} \in S$ を含む軌道空間上の (関数的に) 独立な不変量の数は，

$$\dim S - \dim \mathcal{G} + \dim S_{\boldsymbol{x}}$$

で与えられる．ただし，$S_{\boldsymbol{x}}$ は S の点 $\boldsymbol{x}$ における固定群 ($\boldsymbol{x}$ を $\boldsymbol{x}$ に移す $\mathcal{G}$ の部分群) である． □

$^{\forall}P \in \mathcal{C}$ は，適当な 3 次元射影変換を表現する 4×4 の行列 M を用いて

$$P \;=\; \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} M$$

とかける．これは，対象を固定して視点を変えることと視点を固定して対象を動かすことが等価であることを示す．したがってビジョンにおける不変量の数は，与えられた $\mathcal{P}^3$ の部分集合に対して 3 次元射影変換群が作用するときの

[2] ベクトルはすべて縦ベクトルとする．

[3] $\mathcal{C}$ は一つのカメラで視点を変えて投影像を得る場合の変換だけでなく，視点ごとに異なるカメラで投影像を得る場合の変換も含む．

[4] 具体的には，点の集まりや直線の集まりなどである．

表 1: The number of functionally independent invariants for several configurations under a variety of transformation groups (DOF stands for degree of freedom; * is the case where the stabilizer's dimension is not zero)

Configuration (DOF)	Euclidean (6 DOF)	Affine (12 DOF)	Projective (15 DOF)
5 points (15)	9	3	0
6 points (18)	12	6	3
3 lines (12)	6	0	0
4 lines (16)	10	4	2*
1 line and 4 points (16)	10	4	1
2 lines and 3 points (17)	11	5	2
3 lines and 2 points (18)	12	6	3
quadric (9)	3	0	0
cubic (18)	12	6	3
2 conics on 2 planes (16)	10	4	1

Klein [24] の意味での不変量の数で抑えられる．観測する対象とそれに対する Klein [24] の意味での不変量の数の関係の例を表 1 に示す．"投影による情報の欠落" があるため，Klein [24] の意味での不変量をすべてビジョンの枠組で求めることができるとは限らず，この数はあくまで目安である．

2.2 幾何的 / 代数的不変量

不変量に関する最近の研究成果を，$N = 1$ の場合と $N \geq 2$ の場合に分けて紹介し，幾何と代数の両方の立場から解説する．

2.2.1 1 枚の投影像から得られる不変量 ($N = 1$)

ここでは，1 枚の投影像から得られる不変量について述べる．これまでに導出された不変量とその数の関係を表 2 に示す．ビジョンの分野で長年利用されてきた不変量 [8] は直線上に存在する 4 点から得られる複比のみであり，他の不変量に関心が寄せられ始めたのはごく最近のことである．

不変量は視点に依存せず一定の値をとるため，原理的には 1 枚の投影像からその値を計算することができる．観測対象が 2 次元物体である場合は，たとえそれが 3 次元空間内におかれてあっても，物体と画像面をそれぞれ $\mathcal{P}^2$ に埋め込めば，物体を画像面に投影する変換は 2 次元射影変換となり，"投影による情報の欠落" が生じない．したがって，2 次元

表 2: The number of derived single-view-invariants for specified configurations

Configuration [refs.]	Number of invariants
4 colinear points [8]	1
5 coplanar points [1]	2
5 coplanar lines [31]	2
2 points coplanar with 2 lines [61]	1
2 coplanar conics [13], [14], [37]	2
2 points coplanar with a conic [31]	1
general points [5], [30], [31]	0
6 normal vectors of planes of a trihedron [41], [42]	3
5 lines on 2 planes [51]	1
6 lines on 3 planes [53]	1

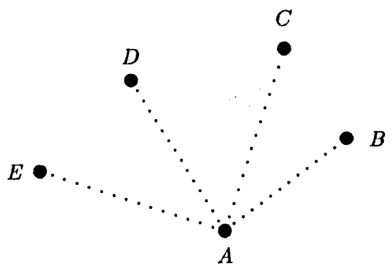

図 1: A cross-ratio of five coplanar points A, B, C, D and E

物体を対象とすると，ビジョンの枠組での不変量の議論と Klein [24] の枠組での不変量の議論は一致する．そこで，まず，2 次元物体を対象とした不変量がビジョンの分野で議論された[5]．その幾つかを解説する．なおここでは，2 次元物体，画像面はともに $\mathcal{P}^2$ に埋め込まれているとする．

- **[平面上の 5 点]** [1] ($\dim S = 10$, $\dim \mathcal{G} = 8$, $\dim S_x = 0$)

  同一平面上の 5 点 (ただし，どの 3 点も同一直線上にない) に対して，次の二つの (関数的に独立な) 不変量が存在する．すなわち，平面上の 5 点の座標を $\boldsymbol{x}_i$ ($i = 1, 2, \ldots, 5$) とすると，

$$I_1 = \frac{\det P_{431} \cdot \det P_{521}}{\det P_{421} \cdot \det P_{531}}, \qquad I_2 = \frac{\det P_{421} \cdot \det P_{532}}{\det P_{432} \cdot \det P_{521}} \tag{2.2}$$

  である．ここに P_{ijk} は，$P_{ijk} = (\boldsymbol{x}_i \ \boldsymbol{x}_j \ \boldsymbol{x}_k)$ なる 3×3 の行列である．勝手な 2 次元射影変換 T によって $\boldsymbol{x}_i$ が $\boldsymbol{x}'_i$ に移されるとすると，$\boldsymbol{x}'_i = \sigma_i T \boldsymbol{x}_i$ (ただし σ_i は実数) が成り立つ．したがって，変換後に対して P'_{ijk} を同様に定義すると，$\det P'_{ijk} = \sigma_i \sigma_j \sigma_k \cdot \det T \cdot \det P_{ijk}$ となるので，I_1, I_2 の値はともに T に依存しない．4 点によって $\mathcal{P}^2$ の標準射影座標枠が定まるので，I_1, I_2 はこの座標枠に対する第5の点の斉次座標の成分の比であると解釈できる．また，この不変量の幾何学的な解釈は，平面上の 5 点から導出される，1 点を通る 4 直線に対する直線の複比となる (図1)．1 点を通る 4 直線の導出には 5 通りあるが，そのうち独立なものは二つである．なお，複比そのものが I_1, I_2 に対応するのでなく，I_1, I_2 は複比の関数として表される (すなわち，複比と関数的に独立でない) という意味である．これは以下にも当てはまる．

- **[平面上の 5 直線]** [31] ($\dim S = 10$, $\dim \mathcal{G} = 8$, $\dim S_x = 0$)

  $\mathcal{P}^2$ 上では点と直線は双対であるから，どの 3 直線も 1 点で交わらない 5 直線に対して，式 (2.2) と同じ不変量が存在する．

- **[平面上の 2 点と 2 直線]** [61] ($\dim S = 8$, $\dim \mathcal{G} = 8$, $\dim S_x = 1$)

  座標が $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ である 2 点と，座標が $\boldsymbol{\ell}_1$, $\boldsymbol{\ell}_2$ である 2 直線に対する不変量は

$$I = \frac{(\boldsymbol{\ell}_1 \cdot \boldsymbol{x}_1)(\boldsymbol{\ell}_2 \cdot \boldsymbol{x}_2)}{(\boldsymbol{\ell}_1 \cdot \boldsymbol{x}_2)(\boldsymbol{\ell}_2 \cdot \boldsymbol{x}_1)}$$

  で与えられる．点に作用する 2 次元射影変換 T によって，直線の座標は $\boldsymbol{\ell}_i$ から $T^{-\mathrm{T}} \boldsymbol{\ell}_i$ に移されるので，I の値は T に依存しない．ただし，$T^{-\mathrm{T}}$ は T^{-1} の転置を表すと約束する．2 直線と 2 点から導出される同一直線上の 4 点 (あるいは，1 点を通る 4 直線) に対する複比が，その幾何学的な解釈となる (図2)．

- **[平面上の二つのコニック]** [13], [14], [37] ($\dim S = 10$, $\dim \mathcal{G} = 8$, $\dim S_x = 0$)

  コニックは，3×3 の対称行列 C を用いて

$$\boldsymbol{x}^{\mathrm{T}} C \boldsymbol{x} = 0$$

  と表せる．C_1, C_2 で表される同一平面上の二つのコニックに対して，次式で表される不変量が存在する:

$$I_1 = \mathrm{trace}[C_1^{-1} C_2](\det C_1 / \det C_2)^{\frac{1}{3}}, \qquad I_2 = \mathrm{trace}[C_2^{-1} C_1](\det C_2 / \det C_1)^{\frac{1}{3}}.$$

[5] Reiss [40] は 2 次元物体を対象とした不変量とその応用を多角的に論じている．

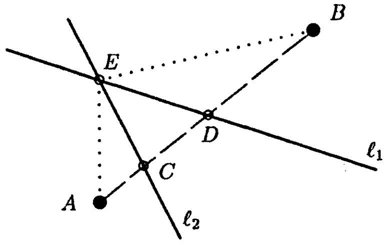

図 2: The cross-ratio of two points, A and B, coplanar with two lines ℓ_1 and ℓ_2 (diag$(1, \alpha, \alpha)$ is the stabilizer of this orbit in the coordinate system where $C(0,0,1)^{\mathrm{T}}$, $D(0,1,0)^{\mathrm{T}}$ and $E(1,0,0)^{\mathrm{T}}$)

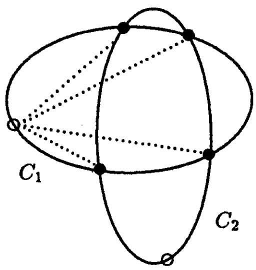

図 3: A cross-ratio of two conis C_1 and C_2

点に作用する 2 次元射影変換 T によって，コニックを表す行列は C_i から $T^{-\mathrm{T}}C_iT^{-1}$ に移されるので，I_1，I_2 の値は T に依存しない．一方，次の性質にもとづいて，二つのコニックが (複素の意味も含めて) 交わる 4 点と各コニック上の任意の 1 点を通る 4 直線に対する複比が，この不変量の幾何学的な解釈である (図3参照)．"一つのコニック上に 4 点が与えられているとする．そのコニック上に 5 番目の点をとり，この点と他の 4 点を通る 4 直線を考える．このとき，この 4 直線に対する複比は 5 番目の点の取り方に依らない．"

このように，2 次元物体を対象とするとビジョンにおける不変量と Klein [24] の意味での不変量の違いを意識しなくてよい．しかし，3 次元物体を対象とすると "投影による情報の欠落" が生じるので，両者の不変量の議論は根本的に異なる．例えば，3 次元空間内の一般の位置にある 6 点に対する 3 次元射影変換群の下での Klein [24] の意味での不変量は 3 個存在するが，3 次元空間内の一般の位置に存在する点の集合に対しては，点を何点選んでも，1 枚の投影像から不変量を導出することはできない [5], [30], [31]．制約のない (一般の)3 次元物体に関して 1 枚の投影像から導出される不変量に対する否定的な結果が得られたのをうけて，3 次元空間内での配置に制約のある点や直線などに対する不変量を求めることが試みられた．そして，1) 三面頂点多面体[6]の 6 個の平面の法線ベクトルに対する 3 個の不変量 [41], [42]，2) 2 枚の平面上の 5 直線に対する 1 個の不変量 [51]，3) 3 枚の平面上の 6 直線に対する 1 個の不変量 [53]，を 1 枚の投影像から導出できることが示された．

– [三面頂点多面体の 6 個の平面の法線ベクトル] [41], [42]

三面頂点多面体の 6 個の平面の $\mathcal{P}^3$ での座標を $\boldsymbol{n}_i$ ($i = 1, 2, \ldots, 6$) とする．このとき，この三面頂点多面体に対する不変量は次式で与えられる:

$$I_1 = \frac{\det N_{3561} \cdot \det N_{3542}}{\det N_{3564} \cdot \det N_{3512}}, \quad I_2 = \frac{\det N_{3562} \cdot \det N_{3142}}{\det N_{3512} \cdot \det N_{3642}}, \quad I_3 = \frac{\det N_{3564} \cdot \det N_{5612}}{\det N_{3561} \cdot \det N_{5642}}. \tag{2.3}$$

ここに $N_{ijk\ell}$ は，$N_{ijk\ell} = (\boldsymbol{n}_i\ \boldsymbol{n}_j\ \boldsymbol{n}_k\ \boldsymbol{n}_\ell)$ なる 4×4 の行列である．ここでは，1) 三面頂点多面体を構成する平面の 3 次元空間内での座標を変数とし，多面体の各頂点の画像面上での座標を係数とする 1 次同次連立方程式を導き，2) "この係数行列のカーネルの次元が 4 であるとき，解は 3 次元射影変換の不定性を除いて一意に表現できる" ことを利用する，という方針で不変量を求めている．すなわち，暗に (3 次元射影変換の不定性を許して)3 次元情報を復元して，3 次元空間内に存在する Klein [24] の意味での不変量を求めている．

[6]全ての頂点が，3 個の面の交わりとして特徴づけられる多面体．

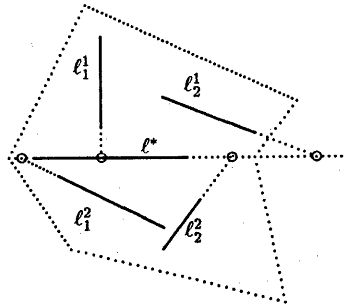

図 4: The cross-ratio of five lines on two planes

- [2 平面上の 5 直線] [51]

  2 平面上の両平面の交線を含む 5 直線に対して，次の不変量が存在する．すなわち，$\mathcal{P}^2$ に埋め込まれた画像面上において，一方の平面上にある 2 直線の座標を ℓ_1^1, ℓ_2^1, 他方の平面上にある 2 直線の座標を ℓ_1^2, ℓ_2^2, 両平面の交線の座標を ℓ^* とすると，

$$I = \frac{\det(\ell_1^1\ \ell^*\ \ell_1^2)\cdot\det(\ell_2^1\ \ell^*\ \ell_2^2)}{\det(\ell_1^1\ \ell^*\ \ell_2^2)\cdot\det(\ell_2^1\ \ell^*\ \ell_1^2)}$$

  である．ここでは，“3 次元空間内では，2 平面上の両平面の交線を含む 3 直線の各解釈面 (視点と直線を含む平面) の法線ベクトルを稜とする平行 6 面体に対して，視点の変化の前後における体積比は各平面上で直線が存在する位置に依存しない” ことを利用して不変量を求めている．この不変量の幾何学的な解釈は，4 直線を 2 平面の交線上にのばすことによって得られる，同一直線上の 4 点に対する複比となる (図4参照).

2.2.2 複数枚の投影像から得られる不変量 ($N \geq 2$)

ここでは，複数枚の投影像から得られる不変量について述べる．複数枚の投影像を用いて不変量を導出する試みは始まったばかりであり，不変量はまだほとんど導出されていないのが現状である．表3にこれまでに導出された不変量を示す．

複数枚の投影像を利用するのは “投影によって欠落した情報” を補うためである．したがって，複数枚の投影像を用いることによって，Klein [24] の意味での不変量をすべてビジョンの枠組で (すなわち投影像から) 求めることができる可能性がある．画像面上の特徴 (点や直線など) の座標からその 3 次元座標を復元できれば，その特徴に対する 3 次元射影変換群の下での不変量を求めることができる．たとえ 3 次元座標を一意に復元できなくても，3 次元射影変換の不定性を除いて一意に復元できれば，やはり不変量を求めることができる．なぜなら，特徴から適当なものを選び，それによって $\mathcal{P}^3$ の座標枠を規定すれば，他の特徴は全てその座標枠による斉次座標で表され，各特徴を表す斉次座標の成分の比は 3 次元射影変換群の下では不変となるからである．

“3 次元空間内の一般の位置に存在する点の集合に対して，キャリブレーションが行われていない 2 枚の投影像の点対応から，それらの点の 3 次元座標を 3 次元射影変換の不定性を除いて一意に復元できる” ことが示された [9], [23]．こ

表 3: The number of derived multiple-view-invariants for specified configurations

Configuration [refs.]	Calibration	Number of used views	Number of invariants
constrained 6 points [1]	fully	2	1
general 6 points [22]	weakly	2	3
general 4 lines [19], [22]	weakly	2	2
2 conics on 2 planes [39]	weakly	2	1
general 6 points [38]	no	3	3

れにより，2枚もしくはそれ以上の投影像の点対応から3次元空間内の点の集合に対する不変量を求めることができる．3次元空間内の一般の位置にある5点によって射影座標枠が一意に決まるので，この座標枠による6番目の点の斉次座標の成分の比は3次元射影変換群の下で不変となる．Quan [38] は，"3次元空間内の一般の位置に存在する6点のキャリブレーションが行われていない3枚の投影像から，3次方程式を解くことによって3個の不変量を求めることができる" ことを示した．ここでは，1) 3次元ユークリッド空間と画像面をそれぞれ $\mathcal{P}^3$ と $\mathcal{P}^2$ に埋め込み，2) 3次元空間内の6点のうちの5点を $\mathcal{P}^3$ の標準射影座標枠にとり，残りの1点をこの座標枠による斉次座標で表し，3) 6点の像のうち4点を $\mathcal{P}^2$ の標準射影座標枠にとり，残りの2点をこの座標枠による斉次座標で表し，4) 投影を通して得られる $\mathcal{P}^3$ から $\mathcal{P}^2$ への対応から，上記斉次座標に関する，$\mathcal{P}^3$ から $\mathcal{P}^2$ への射影変換に依存しない関係 (4変数の2次同次方程式) を導き，5) これを3枚の投影像に対して連立させ，6) それを1変数の3次方程式に帰着させる，という方針で3個の不変量を代数的に求めている (ただしここでは，3組の解が存在する場合があり，その場合，3組のうちどの組が3個の不変量に対応するかは未解決である: 3個の不変量を必ず一意に求めることができるわけではない)．§2.1でみたように，3次元空間内の一般の位置にある6点に対する3次元射影変換群の下での不変量の数は3であるので，この6点に対するKlein [24] の意味での不変量はすべて，ビジョンの枠組で求められたことになる．なお，点対応の場合と違い，キャリブレーションが行われていない投影像の直線対応やコニック対応から，それらの3次元座標を (どういう意味でかも含めて) 復元できるかどうかという問題は未解決である．

キャリブレーションが行われていない場合の，基本行列 (essential matrix) [27] に対応する概念として fundamental 行列 [11] が導入された．そして，2つの異なる視点に対して fundamental 行列が既知であるとき，weakly calibrated であるとよばれる[7]．weakly calibrated な2枚の投影像からの点，直線，あるいはコニック対応による3次元座標の復元は，3次元射影変換の不定性を除いて一意に決まる．すなわち，weakly calibrated を仮定すれば，3次元空間内の点，直線，あるいはコニックに対して，2枚の投影像の点対応，直線対応，あるいはコニック対応から，その配置に対する3次元射影変換群の下での不変量を求めることができる．事実，下記の結果が示された．これらはすべて，1) 点，直線，あるいはコニックの3次元座標を3次元射影変換の不定性を許して復元し，2) そのもとで，その配置に対する3次元射影変換群の下での不変量を求める，という方針をとっている．すなわち，本質的に3次元を復元することで，3次元空間内に存在する Klein [24] の枠組での不変量を求めている．したがって，これらの対象に対する Klein [24] の意味での不変量を全てビジョンの枠組で求めることができるが，ここでの手法は3次元を復元するステップを含むという点で Quan [38] の手法と大きく異なる．また Quan [38] は，epipolar geometry を利用することなく不変量を求めているので，この点でも両手法は異なる．

1. 3次元空間内で一般の位置にある6点の weakly calibrated な投影像2枚から3個の不変量を求めることができる [22].

   - $\mathcal{P}^3$ では点と面は双対であるから，式 (2.3) と同じ不変量が存在する．

2. 3次元空間内で一般の位置にある4直線の weakly calibrated な投影像2枚から2個の不変量を求めることができる [19], [22].

   - 2点を与えれば直線は一意に決まるので，直線を点の組で表す[8]．a_i, b_i を座標にもつ2点を通る直線 i と a_j, b_j を座標にもつ2点を通る直線 j に対して 4×4 の行列 $L_{ij} := (a_i\ b_i\ a_j\ b_j)$ を定義すると，4直線1,2,3,4に対して，

$$I_1 = \frac{\det L_{12} \cdot \det L_{34}}{\det L_{13} \cdot \det L_{24}}, \qquad I_2 = \frac{\det L_{12} \cdot \det L_{34}}{\det L_{14} \cdot \det L_{23}}$$

     はともに3次元射影変換群の下での不変量となる．一方，3次元空間内の直線は線織曲面 (ruled surface) 上の2系に分類され，同系の2直線はねじれの位置にあり，異系の2直線は必ず交わる．一般の位置にある3直線によって線織曲面が一意に定まり (この線織曲面上ではこの3直線は同系に属する)，4番目の直線はこの線織曲面と (複素の意味も含めて) 2点で交わる．そして各点に対して，その点を通り，3直線とは異系に属する直線が一意に定まる．この直線は3直線全てと交わるので，この直線上に4点が存在し，これに対する複比が求まる (図5参照)．こうして得られる二つの複比がこの不変量の幾何学的な解釈である．

[7] weakly calibrated であることは epipolar geometry が既知であることと同値である．具体的には，画像面上で他方の視点の像 epipole がどこに存在するかを既知とすることである．epipole の位置を決定するには7点以上の点対応が必要である [11].

[8] 2平面を与えても直線は一意に決まるが，$\mathcal{P}^3$ では点と面は双対であるので，2点を与えることと2平面を与えることは同じである．

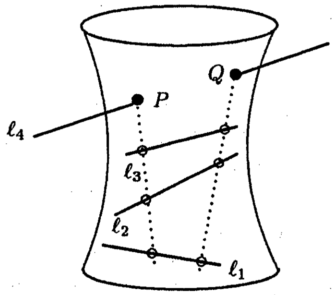

図 5: Two cross-ratios of four lines ℓ_1, ℓ_2, ℓ_3 and ℓ_4

3. 3 次元空間内で 2 枚の平面上にある二つのコニックの weakly calibrated な投影像 2 枚から 1 個の不変量を求めることができる [39].
   - 各コニックは 2 平面の交線と (複素の意味も含めて)2 点を共有するので, 交線上に 4 点が得られる (図6参照). これら 4 点に対する複比がこの不変量の幾何学的な解釈である.

2.3 その他の不変量

ここでは, これまで述べてきた不変量とは異なる範疇に属する不変量について簡単に触れる. 本節で述べる不変量は, 微分不変量 (differential invariants) と準不変量 (quasi-invariants) である.

[微分不変量]

§2.2で取り上げた不変量は, 点や直線, コニックを扱ったものである. これらは, 簡単な (すなわち, 次数の低い) 代数方程式で記述できるものを対象とする, いわば大域的な, 不変量である. しかし, 全ての対象が簡単な代数方程式で記述できるわけではない. 対象を記述する代数方程式の次数が高くなれば, その不変量を導出するのは極めて困難となる. そこで, 対象の, 大域的な性質でなく, 局所的な性質を利用して不変量を求める試みがある. 輪郭がなめらかである対象に対して, 輪郭の微分情報を手がかりとする不変量が微分不変量である. 微分情報は微分をとる位置に依存するので, 微分不変量の値は対象の注目している部分に対してのみ一定となる. 微分不変量は局所情報のみから求められるので次の特長がある: 1) 対象が代数方程式で記述できない場合[9]でも不変量の値を計算することができる; 2) 大域的にみた場合オクルージョンが生じていても, 局所的にオクルージョンが生じていなければ不変量の値を計算することができる. 反面, i) 微分不変量を求めるには一般に高次の微分を扱う必要があり, 輪郭に対して高次の微分可能性が要求される, ii) 微分不変量は局所的であるため, その値を計算するには投影像間の局所的な点対応が必要となる, という欠点がある. 回転軸をもつ 3 次元物体 (すなわち回転面) に対しては, 複接線 (bitangent) や変曲点における接線を扱うことに

[9] この場合でも簡単な代数方程式の当てはめをおこない, 大域的な不変量を利用しようとする研究 [12], [31] もある.

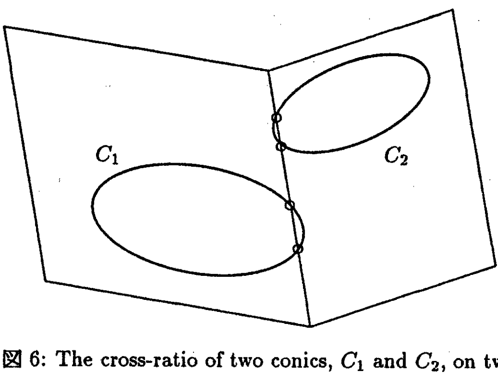

図 6: The cross-ratio of two conics, C_1 and C_2, on two planes

より微分不変量が得られる[15]. なお, 高次の微分の扱いを避けるために, 投影像を標準像に変換し, 標準像に対して微分不変量を議論する報告[43], [60] もある (ただし, これらは2次元物体を扱っている).

[準不変量]

§2.2で述べた不変量は, 3次元空間内で特殊な配置にある点や直線に対するものである. しかし実際の3次元物体を構成する点や直線は必ずしもそのような配置にあるとは限らず, 特殊な配置になければ, 多くの場合不変量は存在しない. そこで, より一般的な場合に対応できるよう, あらゆる視点に対して一定の値を有さずとも, 特徴づけられる集合に含まれる視点に対しては一定の値を有する関数を求める試みがある. 不変量が存在しない点や直線の配置の場合でもこのような関数が存在する可能性がある. この関数は準不変量とよばれ, Binford–Levitt [3] によってその概念が導入された. 準不変量の値が一定となる視点の集合が, 存在しうる視点の大部分を含んでいたり, 現実の状況下で考えられうる視点全てを含んでいたりすると, それは実用上利用価値が高い. 準不変量を求める研究[32] は始まったばかりであり, まだ手探りの状態であるが, 今後どのような見地から研究を進めるべきであるかを説く報告[4] もある.

3 不変量の応用

不変量はビジョンにおけるさまざまな問題と密接に関係している. 本節では, これまで述べてきた不変量のビジョンの課題への応用例として不変量を利用した認識システムをとりあげ, 不変量を利用することによる特長と問題点を指摘する. なお, 本稿では触れないが, 不変量を利用して3次元物体を復元する研究[35], [36], [45], [46], [47] や不変量とキャリブレーションの関係を論じる研究[11], [18], [21], [28] もある.

3.1 認識システムへの応用

物体認識[2] はコンピュータビジョンにおける重要な課題の一つである. ここではモデルに基づく認識に不変量を応用した例[13], [31], [40], [44], [56] を説明する.

モデルに基づく認識における課題は, データベースに予め蓄積された既知の物体のどれが, 与えられた投影像に現れているかを決定することである. データベースには, 物体の輪郭の幾何モデルが蓄積されているのが一般的である. 与えられた投影像内に存在する物体の輪郭がデータベース内のどの幾何モデルの透視変換後の像であるかを決定できたとき, この物体は認識されたとみなされる. 以下では, 不変量を利用しない認識システムと不変量を利用した認識システムを対照させて紹介する. いずれのシステムでも, まず, Step 1: 与えられた投影像から一つの物体を切り出し, そして, Step 2: 一つの物体に属する特徴(点や直線など)のうちデータベース内のモデルと比較すべきものを選出し, Step 3: 選出された特徴をグループに分ける, という手続きが含まれる. なお, ここではデータベース内のモデルの特徴は予めグループに分けられているとする. また, 不変量を利用しないシステムでは, 投影像を得る際にカメラの内部パラメタの決定が行われていることが前提となる.

[不変量を利用しないシステム]

4: Step 3 のグループのうち一つを選ぶ.

  I: データベースのモデルを一つ選ぶ.

    (a): このモデルの特徴のグループを一つ選ぶ.

      a1: 投影像の特徴のグループとモデルの特徴のグループをマッチさせる.

      a2: モデルから投影像への変換 (カメラの外部パラメタ) を決定する.

      a3: この変換に基づいて, モデルの特徴を全て投影像上に投影する.

      a4: モデルの投影像が与えられた投影像に一致するかどうかを判定する.

[不変量を利用したシステム]

0: データベース内のモデルの特徴のグループに対して不変量の値を計算し, それをモデルの属性値として hash table を作成しておく.

4: Step 3 のグループに対して不変量の値を計算する.

5: その値と属性値が一致するモデルを選ぶ.

  I: このモデルを投影像上に投影し, それが与えられた投影像に一致するかどうかを判定する.

不変量を利用しない場合は，与えられた投影像中の選出された特徴の全てのグループ，データベース内の全てのモデル，各モデルの特徴の全てのグループ，に対して繰り返す手続きが含まれるので，データベース内のモデル数が少なくても計算コストは膨大である (モデル数を λ，投影像中の選出された特徴数を i，各モデルの特徴数を m，モデルから投影像への変換を定めるのに要する特徴数を k とすると，物体の候補の生成に必要な計算量は $O(\lambda\, i^k m^k)$ となる)．これに対し不変量を利用すると，1) 認識に要する計算の効率が上がる (不変量の値を計算するのに要する特徴数を ℓ とすると，物体の候補の生成に必要な計算量は $O(i^\ell)$ となる)，という利点に加えて，2) データベースへのアクセスが高速にできる (与えられた投影像内に存在する物体の候補を選ぶ (Step 5) のに要する手間は，データベースの大きさによらない)，3) カメラの内部パラメタを決定する必要がない，4) 物体の一部の特徴だけから不変量の値を計算することができるので，オクルージョンが生じた場合にでも対応できる，などの特長がある．これに対し不変量を利用することによって生じる問題としては，i) 不変量をもつ特徴の配置が限られているためそのグループ化が困難である，ii) 不変量の値は特徴をどの順序で用いるかに依存する[10]ため，一つの特徴のグループからでも得られる値は複数存在する．そしてその数は，不変量の値を計算するのに必要な特徴の数に対して組合せ論的に増加する，iii) 不変量の値を計算するに要する特徴の数は一般に少ないため物体のごく一部の情報だけを比較することになり，投影像を与えた物体の候補を必要以上に選ぶことになりかねない，などが挙げられる．

4 限界と展望

ここでは，ビジョンにおける不変量の研究の現状を踏まえ，不変量をより実用的なものとするために取り組むべき課題について述べる．そして，ビジョンにおける不変量の研究について今後の方向性を探る．

これまでみてきたように，ビジョンにおける不変量の研究は近年盛んに行われている．これにともなって，ビジョンにおけるさまざまな課題に対して不変量の利用価値が唱えられている [31], [58]．事実，数年前に比べて不変量を利用した例の報告は著しく増加している．一見，これまでに得られた不変量に関する結果はただちに利用可能であり，これをそのまま実際の問題に適用できるかのようである．ところが，不変量を実際の問題に応用しようとすると，多くの問題に直面する．以下，そのうちの幾つかを順にあげ，不変量をより実用的なものにするために克服すべき課題について述べる．また，不変量の研究を発展させるための課題についても述べる．

1. 不変量の値の信頼性
   - 理論的には不変量は視点に依存せず一定の値を有する．しかし，実際の投影像には必ずノイズが存在し，不変量の値にばらつきが生じる．理論値からのばらつきがどれほどの範囲内であれば同じ値と見なしてよいかを判断する指針がなければ，実際の投影像から計算した不変量の値を十分に活用できない．しかるに，ノイズに対する不変量の値の振る舞いを論じた報告[11]は皆無である．不変量の値を計算するには大抵 “割り算” が含まれているため，値の信頼性を統計的手法によって理論的に解析することは容易でない．しかしながら，不変量をより実用的なものとするためには不変量の値の信頼性を理論的に論じる必要がある．

2. 不変量の値を計算する特徴の選出
   - 点や直線などの特徴から不変量の値を計算するが，その計算に必要な特徴の数は限られている．しかし，実際の投影像から得られる特徴の数はもっと多く，それらの中から不変量の値を計算するために “適した”(どういう意味で適しているかの基準を定義しなけばならないが) 特徴を選出しなければならない．投影像内に存在する特徴の数に対してこの選出の場合の数は組合せ論的に増加し，“適した” 特徴の選出は困難となる．一方，投影像から得られた特徴が有する情報は同等であるにもかかわらず，一方が選出され他方が選出されないというのは不自然である．また，特徴の選出によって，選出されない特徴から得られる情報を捨ててしまうことになり，投影像から得られる情報すべてを用いないという点で情報の十分な活用が妨げられる．むしろ，必要数以上に特徴がある場合に，それら全てを駆使し，いかに信頼性の高い不変量を導出するかを探るべきである．これは，不変量の値の計算に必要な枚数以上に投影像がある場合の，投影像の選出にも当てはまる．

3. 不変量の値の共有
   - $\mathcal{P}^3$ の部分集合 S を与えると，S に対する不変量の値が一意に決まる．しかし，ある値を決めると，その値をとる集合 S が一意に決まるわけではない．すなわち，不変量は多対 1 の写像である (1 対 1 の写像になる場

[10] ただし，直線上の 4 点に対する複比については，4 点の並び方の順序に依存しない不変量 j-invariant がある．

[11] その多くは特徴にピクセルレベルでノイズを与え，不変量の値の変化を実験的に調べているだけ (例えば，[16]) であり，理論的立場から解析した報告は [29] だけである．

合, 完全な不変量とよばれる). したがって, 投影像から計算された不変量の値によって, その像を与えた物体が一意に決まるわけではない. これは, 不変量が投影像から得られる情報を 1 次元情報に変換する写像である故の必然的な性質であり, 避けられない. しかし, ある不変量に対して, その値を共有する集合がどのように特徴づけられるかを調べることは, その不変量では区別できない物体のクラスを規定することになり, その意義は大きい. さらに, その結果から不変量の性能の比較, ひいては不変量間の階層構造の議論を行うことができる. この問題はまだ論じられていない.

4. 不変量研究の新展開

   – 本稿で述べたように不変量の導出法は, 現状では不変量ごとに異なる. しかし, 一つの手法によって不変量を系統的に導出することができれば, まだ見出されていない不変量の導出は容易になる. 外積代数を発展させた double algebra を利用すると, 点と直線の組合せである配置に限れば, どのような組合せでもその配置に対する不変量を系統的に導出することができる [6], [54]. しかし, 点と直線の組合せでない配置に対する不変量を系統的に導出するような手法は未だ構築されていない. そのような手法を構築する模索はこれからである.

   – これまで得られた不変量はすべて, 2 値化された投影像を扱って導出されている. より多くの情報を用いれば, より多くの, そしてより信頼性の高い不変量を求めることができる可能性がある. 濃淡情報を利用して不変量を論じる試み [59] が報告されているが, この方向への発展はこれからである.

5 CVCV-WG での議論

以下に CVCV-WG のメンバー (敬称略) によって行われた議論の要約を掲載する.

久野 (大阪大): CV で扱っている不変量は数学 (射影幾何学) 的なものであるが, 心理学にも人間の視覚は不変項を知覚しているという理論がある (たとえば, J. J. Gibson: *The Ecological Approach to Visual Perception*, Houghton Mifflin, 1979; 古崎他訳: 生態学的視覚論, サイエンス社, 1985 や宮崎清孝, 上野直樹: 視点, 東京大学出版会, 1985). Gibson は, "それまでの視覚の理論はすべて, ある時点の像が網膜にうつったとしてその像を解釈するというスナップショット的なものであり, それが間違いのもとであった. そうではなく, 視覚は (目や身体を動かすという行動をともなって) 環境の中の不変項を抽出している. そして, 4 種の不変項が仮定される: 1) 変化する照明の下での光学的構造の不変項, 2) 観察点の変化中の光学的構造の不変項, 3) 包囲光の標本抽出を通しての不変項, 4) 包囲光構造の局所的乱れの中の, 包囲光の局所的不変項" と主張している. 1) は人間は照明が変化しても物体の反射特性を認知できるということで, CV でも shape from shading や color constancy (Gibson によれば, 恒常性がどうして存在するのかを議論するのはおかしくて, 恒常的なものをもともと知覚するとのことだが) で検討していることであると思う. 本稿での不変量にもっとも関連するのは 2) の不変項である. 例えば, 人間は目を動かして物体を見たとき, 目の移動に対応した変換後の画像が見えるならば, それは一つのまとまった物体に見える. CRT に正方形を出しそれを透視投影した形のものを順次表示すると, 正方形が 3 次元的に運動しているように見える. このとき, 高さを縮めながら台形にして行くと正方形が倒れて行くように見えるが, 他の変換を施したものを表示すると正方形が変形して行くのを感じる. すなわち, 透視投影に不変なものを物体として知覚している. ところで, CV における不変量に関して次の 2 点が気にかかっているが, それに対してはどのように考えているのか. i) CV における不変量研究とは何か, ii) 複数の視点から導出される不変量は有望だと思うが, 対応点検出が問題となる. 対応点検出にはカメラを少しずつ動かせばよいが, それでは Tomasi and Kanade の factorization method のようなものになるのだろう. しかし, 不変量は indexing に非常に有効なので, こういう形の復元ではなくて, 物体認識にうまく使うという方向がないだろうか.

杉本 (ATR): i) については, 脚注 1 で述べたように, まず "過去の数学の成果を CV にもってくること" が挙げられるが, それにとどまっていてはいけないと考えている. しかし, 現時点で得られている成果は, 数学で発見された不変量を CV にもってくるというレベルにあると思う. 数学でまだやられていないことにチャレンジするのは次のステップであり, 例えば, §4 で述べた濃淡情報を利用した不変量はまさしく新しいチャレンジだと思う. ii) に関しては, 次のように考える: 不変量は "3 次元を復元しないで認識をおこなう" という立場にあると思うが, 複数の投影像を使った不変量はある意味で 3 次元を復元している. したがって, 3 次元を全く復元しないで認識をおこなうのは困難であるという感じがする. しかし, 従来のように 3 次元を完全に復元しなくても, 認識を行えるという示唆がある.

田中 (NTT): i) に関して, 数学における不変量と CV における不変量はどのように違うのか.

杉本 (ATR): 数学での不変量は "変換群に対する不変な性質" を論じるが, CV の枠組で論じなければいけない不変量は "透視変換に対する不変な性質" である. 透視変換は, 群をなさないうえに, 次元の縮退 (3 次元から 2 次元) を生

じさせる．したがってCVでは，“2次元から3次元物体の不変量を求める”という数学では想定されていない定式化の下で不変量を論じなければならない(かといって，数学で得られた結果が全く役に立たないというわけではない)．

久野(大阪大): 数学者が個々の場合に対して不変量を具体的に求めていなかったのであれば，CV研究者が発見してCVに役立てれば，それは十分意義のあることだと思う．structure from motionの初期の研究はCV研究者による再発見だったそうだが，再発見がないとその先に進めないので，それはそれで重要だと思う．しかし，再発見にとどまらずgeometric hashingのように不変量を使った認識法の提案のようなものも重要であるだろう．

杉本(ATR): 数学者が個々の物体について一つ一つ不変量を求めることをしなかったのは，物体を規定する次元の自由度と変換群がもつ次元の自由度という観点から抽象化して不変量を論じたためであると思う．しかし，CVでは現実の画像から抽出できる特徴(エッジやコニック)に対する不変量が必要であるため，個々の場合について不変量を求めているのだと思う．これは，数学は抽象化によってきれいな理論が展開できればよいという立場であるのに対し，CVは現実の画像から目を背けてはいけないという立場である故の違いによるもので，それは仕方のないことだと思う．

井宮(千葉大): 数学史の立場からいえば，射影幾何学の母である画法幾何学ですら本当の意味で完成したのは1930年代のドイツにおいてである．現在CV研究者が再発見しているのは，画法幾何学の定理群である．変動する波面の不変量や，紐や帯の不変量はいまも数理科学の最先端の話題であり，このような数理科学の流れのなかでCVにおける不変量を位置づけることが必要であるだろう．

趙(中央大): 今のところCVでの不変量は数学の枠組での不変量理論を超えていないし，これからもその枠組を活用しないといけないと思う．つまり，Kleinの枠組に当てはまらなければ，何らかの形でそれに当てはまるようにする他はないと思う．一方，Kleinをうち破るためには，Klein流の考え方自体が常識となっている環境が必要であるだろう．

杉本(ATR): Kleinの枠組に帰着するようにして論じる限り，いつまでたってもCVにおける不変量の研究は再発見のレベルを越えられないと思う．Kleinの枠組とは別の土俵をつくり，そこでCVにおける不変量を論じることを目指さなければいけない．これからのCVにおける不変量研究では，いかにCVのカラーを出していくかが重要である．

井宮(千葉大): Radon変換はCT (Computer Tomography)に応用されているが，数学者には偏微分方程式の研究で馴染み深いものであり，何も議論する必要がなかった．しかし，'60－'70年代に，化学，天文学，医療で再発見されて現在に至っている．したがって，再発見でもその目的と利用法によっては意味が全く異なる．CVにおける不変量の研究は，再発見の域をいま出るところだと思う．そのためには，何が必要かを記述法まで含めて検討する必要がある．

志沢(ATR): 現時点で得られているCVの成果は数学で発見されていたものの“再発見”のレベルであるという評価は，数学者の視点に立つものであると思う．CVの科学的側面をみれば，CVの成果が数学的には再発見であることはむしろ当然であると思う．数学者からみて“再発見”であるから，価値が低いということはないと思う．一方，理論物理の問題が新しい数学を作ることがあるように，CVの問題から新しい数学の分野が生まれるとおもしろいと思う．

井宮(千葉大): 不変量の応用はCVではまだ始まったばかりであるが，PR (Pattern Recognition)では古くから—PRとて歴史は30年そこそこであるが—不変特徴量の構成は基本的な話題であり，現在も研究されている．CVにおける不変量とPRにおける不変量との役割の本質的な違い，あるいはまったく同じなのかを明確にすることが今後重要であると思う．

杉本(ATR): CVでは“常にその像をもたらした3次元対象物が意識されている”のに対し，PRでは“はじめにパターンありき”という立場で不変量が論じられていると思う．

井宮(千葉大): 結果的にPRでは不変量は認識のための情報圧縮にしか利用されていない．しかし，CVでは不変量そのものを利用して対象の同定を行うことを目的としていると思う．PRでは不変量をもとにともかく対象を分類すればよいというのに対し，CVでは不変量をもとにその対象の構造を同定する，と考えることができる．これは大きな違いであり，この意味でgeometric hashingもHough transformationもCVであるといえる．

趙(中央大): 不変量に関して，CVとPRをわざわざ区別する必要はないと思う．3次元対象物もパターンとみなすことができるからである．むしろ，PRや数学の中で今まで蓄積されたものから価値のあるものだけを取り出し，足りないものを加えていくという考え方の方が建設的であると思う．

井宮(千葉大): 先に述べたのは，処理の面を強調する立場である．一方，CVとPRにおける不変量を研究の目的によって区別あるいは同一視する立場もあるだろう．CVやPRの研究の目的は，“外界の3次元世界を2次元平面に投影された画像から理解・認識する”である．理解・認識を論じるには，例えば，1)構造の同定や分類を行うレベル，2)記号化や符号化によって構造を言語レベルに変換するレベル，3)言語レベルで同定や分類を行うレベル，といったレベルがあり，その中で不変量研究はどこを担うのかを明確にする必要があるだろう．

栗田(電総研): 不変量を議論する場合，“画像の枚数”による分類もあるが，これに加えて“特徴間対応の必要の有無”による分類もあるだろう．例えば，群論からのアプローチとして以下の文献がある．なお，特徴間の対応を必要としない不変量は“積分不変量”とよばれている(杉本)．

大津 展之: 不変特徴抽出の理論 [I]～[IV] —不変特徴抽出—, 電子通信学会誌, Vol. 69, Nos. 5～8 (1986).

M. Tanaka: On the Representation of the Projected Motion Group in 2+1D, *Pattern Recognition Letters*, 14, 671–678 (1993).

M. Tanaka, T. Kurita and S. Umeyama: Image Understanding via Representation of the Projected Motion Group, *Pattern Recognition Letters*, 15, 993–1001 (1994).

田中 (NTT): 積分不変量を実画像に応用しようとすると, 1) オクルージョンに対処できない, 2) セグメンテーションを正確に行わないと, その値が背景の影響を受ける, という問題がある.

趙 (中央大): 局所的な不変量を使えば点対応が必要となり, 大局的な不変量を用いればオクルージョンに弱い, という実情がある. 1種類の不変量ですべてのことができるわけではなく, 何種類かの不変量と "変量" を組み合わせる必要があると思う. しかし, 表現力が強力でロバストな不変量が見つかれば, それは大いに役立つはずである. なお, 不変量の分類については次の文献がある. H. Wechsler: *Computational Vision*, Academic Press, 1990.

6 おわりに

3次元空間から画像面への透視変換全体を表す C は, 実は, 3次元物体がカメラの前方にある場合と後方にある場合をともに含む. しかし, 現実では3次元物体は常にカメラの前方にあるはずだから, 両者を区別しなければならない (本稿で述べた不変量の議論はこの区別を考慮していない). これを区別して不変量を論じる報告[20] もある. また本稿では, 透視変換を近似しないという立場をとったが, 透視変換をアフィン変換で近似して不変量を論じる立場の研究[25], [26], [47], [48], [49], [50], [56] もある.

以上, ビジョンにおける不変量に関して, Klein [24] の意味での不変量との違いに力点をおいて, 最近の研究成果を代数と幾何の両方の立場から解説した. 本稿で述べたようにビジョンの枠組での不変量は多数存在するが, それらの相互関係については全く論じられていない. これまでに得られた不変量をその特殊な場合として包括するような理論体系の構築が, 不変量の研究を大きく前進させると考えられる. また, 不変量に関する理論的研究もさることながら, 理論的成果とその実用性とのギャップはまだまだ大きいという現状を踏まえ, 不変量をより実用的なものにするための研究も今後に期待したい.

参考文献

[1] E. B. Barrett, P. M. Payton, N. N. Haag and M. H. Brill: General Methods for Determining Projective Invariants in Imagery, *CVGIP: Image Understanding*, 53, 1, 46–65 (1991).

[2] P. J. Besl and R. C. Jain: Three-Dimensional Object Recognition, *ACM Computing Surveys*, 1, 17, 75–145 (1985).

[3] T. O. Binford and T. S. Levitt: Quasi-Invariants: Theory and Exploitation, *Proc. of DARPA Image Understanding Workshop*, 819–829, 1993.

[4] T. O. Binford, D. Kapur and J. L. Mundy: The Relationship Between Invariants and Quasi-Invariants, *Proc. of the 1st ACCV*, 508–511, 1993.

[5] J. B. Burns, R. S. Weiss and E. M. Riseman: View Variation of Point-Set and Line-Segment Features, *IEEE Trans. on PAMI*, 15, 1, 51–68 (1993).

[6] S. Carlsson: Multiple Image Invariants Using Double Algebra, *Proc. of the 2nd DARPA-ESPRIT Workshop on Invariance*, 335–350, 1993.

[7] H. S. M. Coxeter: *Projective Geometry* (2nd edition), Springer-Verlag, New York, U.S.A., 1987.

[8] R. O. Duda and P. E. Hart: *Pattern Classification and Scene Analysis*, Wiley, New York, U.S.A., 1973.

[9] O. D. Faugeras: What can be Seen in Three Dimensions with an Uncalibrated Stereo Rig?, *Proc. of the 2nd ECCV*, 563–578, 1992.

[10] O. Faugeras: *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, Massachusetts, U.S.A., 1993.

[11] O. D. Faugeras, Q.-T. Luong and S. J. Maybank: Camera Self-Calibration: Theory and Experiments, *Proc. of the 2nd ECCV*, 321–334, 1992.

[12] D. A. Forsyth, J. L. Mundy, A. Zisserman and C. M. Brown: Projectively Invariant Representations Using Implicit Algebraic Curves, *Proc. of the 1st ECCV*, 427–436, 1990.

[13] D. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller and C. Rothwell: Invariant Descriptors for 3-D Object Recognition and Pose, *IEEE Trans. on PAMI*, 13, 10, 971-991 (1991).

[14] D. Forsyth, J. L. Mundy, A. Zisserman and C. Rothwell: Invariant Descriptors for 3D Object Recognition and Pose, *Proc. of the 1st DARPA-ESPRIT Workshop on Invariance*, 171–208, 1991.

[15] D. A. Forsyth, J. L. Mundy, A. Zisserman and C. A. Rothwell: Recognizing Rotationally Symmetric Surfaces from Their Outlines, *Proc. of the 2nd ECCV*, 639–647, 1992.

[16] P. Gros and L. Quan: *Projective Invariants for Vision*, Technical Report RT 90 IMAG -15 LIFIA, LIFIA Institute IMAG, Grenoble, France, 1992.

[17] J. Harris: *Algebraic Geometry*, Graduate Texts in Mathematics 133, Springer-Verlag, New York, U.S.A., 1992.

[18] R. Hartley: Camera Calibration Using Line Correspondences, *Proc. of DARPA Image Understanding Workshop*, 361–366, 1993.

[19] R. Hartley: Invariants of Lines in Space, *Proc. of DARPA Image Understanding Workshop*, 737–744, 1993.

[20] R. I. Hartley: Cheirality Invariants, *Proc. of DARPA Image Understanding Workshop*, 745–753, 1993.

[21] R. I. Hartley: Self-Calibration from Multiple Views with a Rotating Camera, *Proc. of the 3rd ECCV*, Vol. 1, 471–478, 1994.

[22] R. I. Hartley: Projective Reconstruction and Invariants from Multiple Images, *IEEE Trans. on PAMI*, 16, 10, 1036–1041 (1994).

[23] R. Hartley, R. Gupta and T. Chang: Stereo from Uncalibrated Cameras, *Proc. of CVPR*, 761–764, 1992.

[24] F. Klein: Vergleichende Betrachtungen über neuere geometrische Forschungen, *Math. Ann.*, 43, 63–100 (1893).

[25] J. J. Koenderink and A. J. Van Doorn: Affine Structure from Motion, *J. of the Optical Soc. of America*, A-8, 2, 377–385 (1991).

[26] Y. Lamdan, J. T. Schwartz and H. J. Wolfson: Object Recognition by Affine Invariant Matching, *Proc. of CVPR*, 335–344, 1988.

[27] H. C. Longuet-Higgins: A Computer Algorithm for Reconstructing a Scene from Two Projections, *Nature*, 293, 5828, 133–135 (1981).

[28] S. J. Maybank and O. D. Faugeras: A Theory of Self-Calibration for a Moving Camera, *Int. J. of Computer Vision*, 8, 2, 123–151 (1992).

[29] L. Morin, P. Brand and R. Mohr: Indexing with Projective Invariants, *Proc. of Int. Workshop on Syntactic and Structural Pattern Recognition*, Haifa, Israel, 1994.

[30] Y. Moses and S. Ullman: Limitations of Non Model-Based Recognition Systems, *Proc. of the 2nd ECCV*, 820–828, 1992.

[31] J. L. Mundy and A. Zisserman eds.: *Geometric Invariance in Computer Vision*, MIT Press, Cambridge, Massachusetts, U.S.A., 1992.

[32] R. Navatia and M. Zerroug: Quasi-Invariant Properties and 3-D Shape Recovery of Non-Straight, Non-Constant Generalized Cylinders, *Prof. of DARPA Image Understanding Workshop*, 725–735, 1993.

[33] A. N. Parshin and I. R. Shafarevich eds.: *Algebraic Geometry IV*, Encyclopedia of Mathematical Sciences Vol. 55, Springer-Verlag, Berlin, Germany, 1994.

[34] T. Poggio, V. Torre and C. Koch: Computational Vision and Regularization Theory, *Nature*, 317, 6035, 314–319 (1985).

[35] J. Ponce, T. A. Cass and D. H. Marimont: *Relative Stereo and Motion Reconstruction*, Technical Report UIUC-BI-AI-RCV-93-07, Beckman Institute, Univ. of Illinois, Illinois, U.S.A., 1993.

[36] J. Ponce, D. H. Marimont and T. A. Cass: Analytical Methods from Uncalibrated Stereo and Motion Reconstruction, *Proc. of the 3rd ECCV*, Vol. 1, 463–470, 1994.

[37] L. Quan, P. Gros and R. Mohr: Invariants of a Pair of Conics Revisited, *image and vision computing*, 10, 5, 319–323 (1992).

[38] L. Quan: Invariants of 6 Points from 3 Uncalibrated Images, *Proc. of the 3rd ECCV*, Vol. 2, 459–470, 1994; *IEEE Trans. on PAMI*, 17, 1, 34–46 (1995).

[39] L. Quan: *Invariant of a Pair of Non-coplanar Conics in Space: Definition, Geometric Interpretation and Computation*, Technical Report, LIFIA Institute IMAG, Grenoble, France, 1994.

[40] T. H. Reiss: *Recognizing Planar Objects Using Invariant Image Features*, Lecture Notes in Computer Science 676, Springer-Verlag, Berlin, Germany, 1993.

[41] C. A. Rothwell, D. A. Forsyth, A. Zisserman and J. L. Mundy: *Extracting Projective Information from Single Views of 3D Point Sets*, TR OUEL 1973/93, Dept. of Engineering Science, Oxford Univ., Oxford, U.K., 1993.

[42] C. A. Rothwell, D. A. Forsyth, A. Zisserman and J. L. Mundy: Extracting Projective Structure from Single Perspective Views of 3D Point Sets, *Proc. of the 4th ICCV*, 573–582, 1993.

[43] C. A. Rothwell, A. Zisserman, D. A. Forsyth and J. L. Mundy: Canonical Frames for Planar Object Recognition, *Proc. of the 2nd ECCV*, 757–772, 1992.

[44] C. A. Rothwell, A. Zisserman, J. L. Mundy and D. A. Forsyth: Efficient Model Library Access by Projectively Invariant Indexing Functions, *Proc. of CVPR*, 109–114, 1992.

[45] A. Shashua: Projective Invariant and Structure from Two Perspective/Orthographic Views: Motion and Recognition, *Proc. of DARPA Image Understanding Workshop*, 767–776, 1993.

[46] A. Shashua: Projective Depth: A Geometric Invariant for 3D Reconstruction from Two Perspective/Orthographic Views and for Visual Recognition, *Proc. of the 4th ICCV*, 583–590, 1993.

[47] A. Shashua: On Geometric and Algebraic Aspects of 3D Affine and Projective Structures from Perspective 2D Views, *Proc. of the 2nd DARPA-ESPRIT Workshop on Invariance*, 87–112, 1993.

[48] A. Shashua: *Algebraic Functions for Recognition*, Technical Report A.I. Memo 1452, MIT, Massachusetts, U.S.A., 1994.

[49] A. Shashua: Trilinearity in Visual Recognition by Alignment, *Proc. of the 3rd ECCV*, Vol. 1, 479–484, 1994.

[50] A. Shashua and N. Navab: Relative Affine Structure: Theory and Application to 3D Reconstruction from Perspective Views, *Proc. of CVPR*, 483–489, 1994.

[51] A. Sugimoto: *Projective Invariant of Lines on Adjacent Planar Regions in a Single View*, ATR Technical Report TR-H-034, ATR, Kyoto, Japan, 1993.

[52] A. Sugimoto: *Object Recognition by Combining Paraperspective Images*, ATR Technical Report TR-H-072, ATR, Kyoto, Japan, 1994.

[53] A. Sugimoto: Geometric Invariant of Noncoplanar Lines in a Single View, *Proc. of the 12th Int. Conf. on Pattern Recognition*, Vol. 1, 190–195, 1994.

[54] L. Svensson: *On the Use of the Double Algebra in Computer Vision*, Technical Report TRITA-NA-P9310 (CVAP122), Royal Institute of Technology, Stockholm, Sweden, 1993.

[55] S. Ullman and R. Basri: Recognition by Linear Combinations of Models, *IEEE Trans. on PAMI*, 13, 10, 992–1006 (1991).

[56] D. Weinshall: Model based Invariants for 3-D Vision, *Int. J. of Computer Vision*, 10, 1 27–42 (1993).

[57] I. Weiss: Projective Invariants of Shapes, *Proc. of DARPA Image Understanding Workshop*, 1125–1134, 1988.

[58] I. Weiss: Geometric Invariants and Object Recognition, *Int. J. of Computer Vision*, 10, 3, 207–231 (1993).

[59] I. Weiss: Invariants for Recovering Shape from Shading, *Proc. of the 2nd DARPA-ESPRIT Workshop on Invariance*, 365–376, 1993.

[60] I. Weiss: Noise-Resistant Invariants of Curves, *IEEE Trans. on PAMI*, 15, 9, 943–948 (1993).

[61] I. Weiss, P. Meer and S. M. Dunn: Robustness of Algebraic Invariants, *Proc. of the 1st DARPA-ESPRIT Workshop on Invariance*, 345–358, 1991.