# Fenomenologia e Cognitive Vision

dott. Carmelo Calì
Dipartimento Fieri-Aglaia, Università degli Studi di Palermo
carmelo.cali@unipa.it

## 0. Introduzione

Vernon (2006) suggerisce di individuare un consenso in *Cognitive Vision* nella ricerca di capacità e attributi non-funzionali per un sistema robusto e adattivo.
Questo contributo presenta una teoria informale di capacità percettive e motorie corrispondenti alle proprietà spaziali fenomeniche degli oggetti, vale a dire ai percetti la cui struttura è studiata in fenomenologia della percezione. Infine, se ne illustrano alcune implicazioni rispetto alle teorie cognitiviste e sensomotorie.

## 1. Spazi delle apparenze e dei movimenti.

La teoria è basata su Husserl (1973) e Smith (1995) e impiega una nozione di spazio astratto ispirata a Gärdenfors (2000) considerata da Chella, Frixione, Gaglio (2003) utile per analizzare, formalizzare e implementare l'*anchoring* percettivo.
Uno spazio concettuale è uno spazio metrico $n$-dimensionale di qualità in cui $X_i$ è l'insieme dei valori dell'$i$-esima dimensione qualitativa per $1 \leq i \leq n$ con $n \in N$. Sia **A** lo spazio astratto delle apparenze dell'ambiente dell'agente. Esso rappresenta le apparenze come quantità osservabili variabili in modo continuo. **A** è lo spazio delle proprietà spaziali fenomeniche degli oggetti, in cui le apparenze $a \in \mathbf{A}$ con $a\{k_i, e_{i,}, q_i\}$ sono individuate dai valori delle seguenti dimensioni:
– $K$: campo visivo;
– $E$: estensione;
– $Q$: qualità.
I componenti di $K$, $E$, $Q$, collegati da relazioni specificabili, sono rilevanti per far emergere le proprietà spaziali degli oggetti. Tali relazioni, qualificate dalle trasformazioni esposte in seguito, consentirebbero di individuare la distanza tra due o più apparenze candidate a essere raggruppate come aspetti di uno o più oggetti.
$K$ è semplicemente connesso, in sé congruente, finito e delimitato. $E$ designa estensioni di $a \in \mathbf{A}$ con proprietà geometriche. Tra $K$ e $E$ vale la relazione: $E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ riempiono $K^k\{(k_{i,}, k_j, \ldots, k_n)\}$ e ogni alterazione degli uni comporta una variazione degli altri e viceversa.
$E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ sono reciprocamente delimitati da margini da cui sono scomponibili in parti che a scale inferiori, entro una certa soglia, appaiono autonome.
$K$ ne fornisce un ordinamento stabile per posizioni. Per ogni scomposizione disgiunta del campo: i punti forniscono luoghi per ogni $e_{(i)}$ che si distingue per posizione da $e_{(j,\ldots,n) \neq i}$; le parti ottenute sono localizzate fino ai margini ai limiti di discriminabilità.
$K$ corrisponde a un continuo bidimensionale: per ogni scomposizione di $E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ si ottengono margini estesi ulteriormente scomponibili che si delimitano reciprocamente tramite punti limite.
$Q$ designa qualità (colore, superficie) diffuse su $E$, localizzate in $K$. Ciascuna variazione di $Q$ richiede $E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ e ogni $e \in E$ è dotato di elementi di $Q$. Quindi, ogni $q \in Q$ appare in $K$.
Le qualità contrassegnano i margini di $E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ segregando $a_{j(k,e,q)}$. $E^k\{(e_{i,}, e_j, \ldots, e_n)\}$ distinti individuano $K^k\{(k_{i,}, k_j, \ldots, k_n)\}$ che ordinano localmente colori e superfici. Questa funzione dipende dalle discontinuità di $Q$. Gli elementi di $Q$ si unificano o segregano per discontinuità in funzione del sotto-spazio qualitativo cui appartengono.
Siano $Q^{ki}\{(q^i_{i,}, q^i_j, \ldots, q^i_n)\}$ sotto-spazi di $Q$: $q_{(i,j,\ldots,n)}$ designano differenti qualità (colore, tessitura); $q^i$ il grado di qualità dipendenti (per il colore: tonalità, luminosità, saturazione). Ogni $q^i_{(i,j,\ldots,n)}$ è funzione di elementi definiti

di $E$. Un elemento di $E$ è regolare se la composizione $[(e_{(i,j,..,n)})*(q^i_{(i,j,\dots,n)})]$ – per semplicità $E(Q)$ – è continuo in un intorno di $e_{(i,j,..,n)}$. Siano $O$ l'insieme aperto di punti regolari di $E$ contenente un intorno di ciascuno, e $C$ l'insieme complementare. $C$ è l'insieme chiuso dei punti non regolari perché segnati dalle discontinuità $q^i_{(i,j,\dots,n)}$ che distinguono $e_{(i)}$ da $e_{(j,..,n)\neq i}$ in $K$.
Gli elementi di $Q$ sono ripetibili, mentre i punti di $K$ sono riempiti da un $q^i_{(i,j,\dots,n)}$ alla volta.
L'ordinamento di $K$ è quindi stabile:
- $Q$ varia senza mutare posizione e viceversa;
- l'inversione della traslazione di $a_{j(k,e,q)}$ riconduce allo stesso punto;
- la rotazione di $a_{j(k,e,q)}$ ne muta la distanza relativa, non localizzazione;
- continuità e discontinuità di $Q$ richiedono la variabilità continua di $E$ e $K$.
Sia **M** lo spazio concettuale dei movimenti dell'agente. **M** ha sotto-spazi definiti da capacità oculomotorie, cefalomotorie, cefalo-tronco-motorie, locomotorie. Ogni sotto-spazio è una varietà $n$-dimensionale con sistemi di coordinate le cui trasformazioni individuano gli osservabili dell'agente nell'ambiente.
**A** e **M** sono coordinati funzionalmente. Gli $m\in\mathbf{M}$ fanno cogliere gli $a\in\mathbf{A}$ ma non ne determinano i componenti, né ciascun $M^k\{(m^d_{i,},m^d_{j},\dots,m^d_{n-1})\}$ – con $d$ che designa la dimensione di $m\in\mathbf{M}$ data la trasformazione delle coordinate ammessa nel sotto-spazio – è necessario per un $a\in\mathbf{A}$ che ne è quindi separabile.
Ciascun $M^k\{(m^d_{i,},m^d_{j},\dots,m^d_{n})\}$ ha associato un indice temporale $t^s_{(0,1,\dots n)}$ a scale diverse.
La coordinazione tra **A** e **M** è efficace perché per un indice $t^s_{(0,1,\dots n)}$:
- $m^d_{(i)}$ è selezionato tra gli $m\in\mathbf{M}$ e estrae $a_{j(k,e,q)}$ tra gli $a\in\mathbf{A}$ possibili per $m^d_{(j,..,n)\neq i}$;
- $m^d_{(i)}$ può essere abbandonato per $m^d_{(j)}$.
Quindi, una serie di apparenze coordinate con $M^k\{(m^d_{i,},m^d_{j},\dots,m^d_{n-1})\}$ designerà una linea in **A**. Così l'agente sa anticipare le apparenze, invertirle, pianificare (comportamento attuale, modulazione, deliberazione).
La coordinazione funzionale tra $m\in\mathbf{M}$ e $E(Q)$ si basa sull'associazione fissa tra ciascun $m\in\mathbf{M}$ e l'intero $K$: per $M^k\{(m^d_{i,},m^d_{j},\dots,m^d_{n})\}$ vale sempre l'ordinamento $K^k\{(k_{i,},k_j,\dots,k_n)\}$ tale che ogni $m^d_{(i)}$ ha valore per l'agente rispetto a $K^k\{(k_{i,},k_j,\dots,k_{n-1})\}$ e ogni insieme di $K$ è accessibile tramite $m^d_{(i,j,..,n-1)}$.

2. Sistemi e trasformazioni.

I sistemi di coordinate di **M** in cui **A** varia sono i seguenti.
Per il sistema oculomotorio si ha un piano delimitato con punto zero (posizione fondamentale degli occhi) e assi alto-basso, destra-sinistra con le direzioni delle possibili variazioni di $a\in\mathbf{A}$.
Il movimento oculare amplia $K$ con variazioni cicliche di $a_{j(k,e,q)}$ preservandone le caratteristiche. Infatti, $K$ si modifica gradualmente solo dai bordi e in direzione centro-periferia per $a_{j(k,e,q)}$ che ne entra e fuoriesce. Data la stabilità delle posizioni per gli $E(Q)$ che le riempiono, le apparenze acquisiscono invarianza rispetto a $K$.
Divengono così accessibili alcune proprietà. Limitandosi alla distanza apparente: se una coppia $[(a_{i(k,e,q)}),(a_{j(k,e,q)})]$ scivola dal centro alla periferia per spostamento di $K$ modulo $m^d_{(i)}$, la distanza è preservata.
Ciò fa apparire coppie di punti come una configurazione; configurazioni come una serie $\mathrm{R}_i\{(a_1, a_2, \dots, a_n^{\,i})\in(\mathrm{A}/\sim)_n^{\,i}\}$; una serie come apparenze di uno o oggetti disgiunti tale che all'ordine $K$ corrisponda l'ordine delle loro parti. Gli oggetti compaiono come varietà di dimensione superiore le cui parti si mostrano tramite $a\in\mathbf{A}$.
Per il sistema cefalomotorio, si ha uno spazio cilindrico finito, illimitato per la linea chiusa in direzione sinistra-destra della rotazione della testa; delimitato per la linea verticale alto-basso dell'elevazione della testa, con punto zero fissato dalle posizioni fondamentale degli occhi e normale della testa.
Aggiungendovi le inclinazioni del tronco, si hanno i seguenti $m\in\mathbf{M}$:
- rotazione di $K$: $E(Q)$ preserva la corrispondenza biunivoca;

- dilatazione: $E(Q)$ si espande e contrae preservando la similitudine e apparendo come un oggetto a lontananze diverse. La preservazione di $E(Q)$ separa $a_{j(k,e,q)}$ da ciò che non si dilata;
- spostamento per traslazione o rotazione con occlusione: $a_{i(k,e,q)}$ scivola gradualmente e copre parzialmente o totalmente $a_{j(k,e,q)}$. La distruzione-ricostruzione dell'$E(Q)$ occluso consente di completarlo percettivamente; data la non ripetibilità di $K$, ciò fa apparire una prima forma di profondità.
- deformazioni: $E(Q)$ si deforma in maniera uniforme o meno. Nel secondo caso, la variazione informa sull'appartenenza di $\mathrm{R}_i\{(a_1, a_2, \ldots, a_n^i) \in (\mathrm{A}/\sim)_n^i\}$ allo stesso oggetto che non comprende ciò che non varia o varia diversamente.
- rotazione e deformazione intorno agli assi dell'oggetto: variazione ciclica per cui i gruppi $E(Q)$ che si appartengono alterano inclinazione e orientamento in una sequenza in cui si occludono in funzione della direzione di variazione per riapparire alla fine del ciclo. A differenza dello spostamento, il gruppo non preserva la trasformazione lineare delle coordinate e i punti modificano l'orientamento e si deformano ma univocamente. Diversamente dalla dilatazione, le cui variazioni sono undimensionali e riguardano solo il lato di un oggetto, ora le trasformazioni formano una varietà bidimensionale tale da formare una superficie chiusa con indici di tridimensionalità.

Per il sistema locomotorio, si hanno: (i) uno spazio delimitato con punto zero mobile a partire dal quale è vicino ciò la cui serie di apparenze si modifica con $M^k\{(m^d{}_{i,}, m^d{}_{j}, \ldots, m^d{}_{n-1})\}$, altrimenti è lontano; (ii) uno spazio infinito con tutti i punti equivalenti: ciascuno può divenire un punto zero per $m \in \mathbf{M}$ periodici e invertibili. L'orizzonte si sposta continuamente, preservandosi come caso limite. Le trasformazioni di coordinate fanno apparire una tridimensionalità completa degli oggetti e, compensando lo spostamento dell'orizzonte, dello spazio.

## 3. Conclusione.

La teoria è generalizzabile e soddisfa la richiesta di Vernon (2006) di un *framework* comune in *Cognitive Vision*. È calibrata sulla percezione umana dello spazio ordinario, quindi è integrabile con le evidenze della fenomenologia sperimentale sulle strutture del mondo visivo potenzialmente utili per la costruzione di agenti efficienti (Kanizsa, 1980; Bozzi, 1968, 1989; Vicario, 1973). D'altro canto, **A** e **M** sono concettualizzati come spazi astratti. Quindi, se li si intende come modelli, la teoria proposta può esserne una delle interpretazioni che li soddisfa ma è sostituibile da altre specificabili per agenti artificiali.

Come l'approccio cognitivista, i cui caratteri essenziali sono esposti da Vernon (2006), Granlund (1999) e tradizionalmente rappresentati da Marr (1982), la teoria mantiene il ruolo delle invarianze nell'informazione sulle regolarità ambientali ma senza farle dipendere da un modello a priori. Esse sono estratte dall'agente dalle modificazioni indotte nell'ambiente.

Come nell'approccio sensomotorio (Philipona et al., 2003), **A** è coordinato a **M** ma non vi è ridotto. L'accoppiamento funzionale delle loro dimensioni seleziona i percetti efficaci tra quelli possibili. Si prevede, dunque, che i percetti non siano usati solo indirettamente come consequenze osservabili delle azioni, bensì come strutture di apparenze per sintonizzarsi all'ordine dell'ambiente, senza tuttavia impegnarsi su un modello pre-determinato del reale che è invece ricostruito dall'interno dell'esperienza percettiva.

Rispetto alle teorie ibride (Granlund, 2006), il ciclo percezione-azione della teoria non riduce la percezione a funzione dell'azione. È ipotizzabile anzi che perciò essa trovi applicazione nel testarne le assunzioni per la scelta delle tecniche di derivazione dei parametri dello spazio percettivo utili per la modulazione e pianificazione del comportamento.

Bibliografia


Bozzi, P. (1968) Unità, identità, causalità. Una introduzione allo studio della percezione, Cappelli, Bologna.

Bozzi, P. (1989) Sulla preistoria della fisica ingenua, Sistemi Intelligenti, 1(1), 61-74.

Chella, A., Frixione, M., Gaglio, S. (2003) Conceptual spaces for anchoring, Robotics and Autonomous Systems, 43(2-3),193–195.

Gärdenfors, P. (2000) Conceptual Spaces, MIT Press, Bradford Books, Cambridge (MA).

Granlund, G.H. (1999) The Complexity of Vision, Signal Processing, 74, 101-126.

Granlund, G.H. (2006) Organization of Architectures for Cognitive Vision Systems. In Christensen, H.I., Nagel, H.H. (eds.) Cognitive Vision Systems, LCNS 3948, 37-55, Springer, Berlin.

Husserl, E. (1973) Ding und Raum (1907), Nijhoff, The Hague.

Kanizsa, G (1980) Grammatica del vedere, Il Mulino, Bologna.

Marr, D. (1982) Vision: A computational investigation into the human representation and processing of visual information, W.H. Freeman, San Francisco.

Philipona, D., O'Regan, J.K., Nadal J.P. (2003) Is There something out there? Infering space from sensorimotor dependencies, Neural Computation, 15(9): 2029-2049.

Smith, B. (1995) The Structure of the Common Sense World, Acta Philosophica Fennica, 58, 290-317.

Vernon, D. (2006) The Space of Cognitive Vision. In Christensen, H.I., Nagel, H.H. (eds.) Cognitive Vision Systems, LCNS 3948, 7-24, Springer, Berlin.

Vicario G.B. (1973) Tempo psicologico e eventi, Giunti-Barbera, Firenze.